

**STAGING KNOWLEDGE AND EXPERIENCE:  
HOW TO TAKE ADVANTAGE OF REPRESENTATIONAL  
TECHNOLOGIES IN EDUCATION AND TRAINING?**



August 28-31, 2012

Laboratoire des Sciences de l'Education  
Université Pierre-Mendès-France, Grenoble, France

Editors: Erica de Vries & Katharina Scheiter



## PREFACE

It is a great pleasure to welcome you at the 2012 EARLI Special Interest Group 2 meeting hosted by the Laboratory of Educational Sciences of the Pierre-Mendès-France University in Grenoble, France. The EARLI SIG 2 is one of the twenty-two special interest groups of the European Association of Learning and Instruction. It focuses on how learning is influenced by the representation of the learning material, such as by way of text and pictures. However, as there is an explosion of new representations by the introduction of graphical interfaces, members of the SIG now consider multiple forms of representation, including, but not limited to, texts, pictures, graphs, diagrams, concept maps, animations, equations, virtual reality, information and scientific visualization, haptics, multimedia, hypermedia, and simulations. The SIG meets every two years, in alternation with the main EARLI conference, in order to establish continuity in collaborations of its members.

The theme of the Grenoble 2012 meeting is “Staging knowledge and experience: How to take advantage of representational technologies in education and training?” The term “staging” is used here to suggest the analogy with the “mise-en-scène”, i.e. the design of the visual aspects, in theatre and film productions. In effect, instructional designers, teachers, and trainers, but also learners themselves, can be thought of as “directors” striving to find an appropriate arrangement – in terms of composition, sets, props, actors, costumes, and lighting – of knowledge and experience to be acquired. As mentioned in the SIG 2 mission statement, this nowadays involves using a whole range of possibilities for external representation, not just texts and pictures, as a kind of semiotic technology for expressing knowledge and displaying skills.

The meeting features three keynotes, a JURE pre-conference workshop, eight paper sessions and a poster session. We received a total of 75 paper and poster submissions by 176 authors from 19 different countries. All submissions were reviewed by the scientific committee of 20 members from 9 different countries. 37 papers and 36 posters were accepted. This volume contains all presented papers and posters. The high number of multiple-author, multiple country contributions shows that the SIG is in good shape and the bi-annual meeting fulfils its function.

We thank the European Association for Research on Learning and Instruction, the Scientific Board of the Pierre-Mendès-France University, the Grenoble Alpes Metropole, the City of Grenoble, and the Rhone-Alpes Region for financial support. Finally, we thank in particular the local committee and members of the Laboratory of Educational Sciences for organizing what we hope will be a very fruitful meeting!

Bienvenue et bon colloque.

Erica de Vries  
Katharina Scheiter

# COMMITTEES

## **Chairs**

Erica de Vries, LSE, Pierre-Mendès-France University, Grenoble, France  
Katharina Scheiter, Knowledge Media Research Center, Tübingen, Germany

## **Program Committee**

Shaaron Ainsworth, University of Nottingham, United Kingdom  
Marije van Amelsvoort, Tilburg University, The Netherlands  
Mireille Bétrancourt, University of Geneva, Switzerland  
Maryse Bianco, LSE, Pierre-Mendès-France University, Grenoble, France  
Raquel Cerdán, University of Valencia, Spain  
Jean-Michel Boucheix, LEAD-CNRS, Dijon, France  
Ivar Bråten, University of Oslo, Norway  
Laurent Lima, LSE, Pierre-Mendès-France University, Grenoble, France  
Ric Lowe, Curtin University, Australia  
Alfons Maes, Tilburg University, The Netherlands  
Peggy van Meter, Penn State University, USA  
Jean-François Rouet, LACO-CNRS, Poitiers, France  
Ladislao Salmeron, University of Valencia, Spain  
Tina Seufert, University of Ulm, Germany  
Wolfgang Schnotz, University of Koblenz-Landau, Germany  
Neil Schwartz, California State University at Chico, USA  
Huib Tabbers, Erasmus University Rotterdam, The Netherlands  
André Tricot, CLLE-CNRS, Toulouse, France

## **Local Committee**

Erica de Vries (chair)  
Muhammad Ashraf  
Jonathan Groff (SIG 2 Junior Coordinator, IFSTTAR Lyon)  
Marine Hascoët  
Gwenaëlle Joët  
Sandrine Lescure  
Laurent Lima  
Céline Meurger  
Capucine Mezeix  
Cécile Nurra  
Jérémy Pouille

# TABLE OF CONTENTS

## Keynotes

- The early development of text comprehension skills: emotional inferences are relevant cues..... 9  
*Nathalie Blanc*
- Of metaphors, maps, and methods: cognitively inspired and perceptually salient graphic displays..... 10  
*Sara Irina Fabrikant*
- What kind of graphic is this? A framework for delineating the graphics in text-graphic research..... 11  
*Neil H. Schwartz, Robert W. Danielson*

## JURE Pre-Conference Workshop

- An eye tracking approach to analyze dynamic representations..... 12  
*Jean-Michel Boucheix, Jonathan Groff*

## Papers

- Effects of constructing concept maps while navigating in a hypertext ..... 13  
*Franck Amadieu, Julie Lemarié, Aline Chevalier, Julien Cegarra, Ladislao Salmerón, Stéphane Blat*
- Visualizing football statistics: performance & preference ..... 16  
*Marije van Amelsvoort, Hans Westerbeek*
- Metacognition in reading comprehension: descriptive study ..... 19  
*Olga Arias, Raquel Fidalgo*
- Examining the integration of text and pictures..... 22  
*Jana Arndt, Anne Schüller, Katharina Scheiter*
- Categorization of graphical representations ..... 25  
*Muhammad Ashraf*
- Mere presence, object orientation and perspective taking..... 28  
*Adriana Baltaretu, Alfons Maes, Carel van Wijk*
- A comparison of different levels of interactions when using the isolated-elements strategy..... 31  
*Dominique Bellec, André Tricot, Paul Ayres*
- Using 3D animation for learning functional anatomy ..... 34  
*Sandra Berney, Gaëlle Molinari, Nady Hoyek, Mireille Bétrancourt*
- Graphicacy: Do readers of science textbooks need it? ..... 37  
*Mireille Bétrancourt, Shaaron Ainsworth, Erica de Vries, Jean-Michel Boucheix & Richard K. Lowe*
- The influence of annotation in graphical organizers ..... 40  
*Eniko Bezdan, Liesbeth Kester, Paul Kirschner*
- Paired graphics: an exploratory study of graphicacy ..... 43  
*Jean-Michel Boucheix, Richard K. Lowe, Shaaron Ainsworth, Mireille Bétrancourt, Erica de Vries*
- Reading their way into science: students' strategies for comprehending research articles..... 46  
*Isabel Braun, Matthias Nückles*
- Limits and potentials of bottomup processing of multiple external chemical representations..... 49  
*David Corradi, Jan Elen, Geraldine Clarebout*
- Automated analysis of pupils' self-explanations of a narrative text ..... 52  
*Philippe Dessus, Maryse Bianco, Aurélie Nardy, Françoise Toffa, Mihai Dascalu, Stefan Trausan-Matu*

Students' eye movements when solving mathematical word problems together with illustrations .....	55
<i>Tinne Dewolf, Wim Van Dooren, Frouke Hermens, Lieven Verschaffel</i>	
De-symbolization in learning a new symbolic system .....	58
<i>Gilles Dieumegard</i>	
How a picture can scaffold comprehension of text .....	61
<i>Alexander Eitel, Katharina Scheiter, Anne Schüller, Marcus Nyström, Kenneth Holmqvist</i>	
Graphs in science education – german biology teachers' beliefs, instructional practices, and related student outcomes .....	64
<i>Carolyn Enzengmüller, Claudia Nerdel, Helmut Precht</i>	
Signals in expository prose: is the ability to recognize and interpret them specific of reading comprehension? .....	67
<i>Carlos Ferreira, Hector Garcia Rodicio, Emilio Sanchez</i>	
On-line assessment of students' global reading strategies through triple task technique.....	70
<i>Raquel Fidalgo, Olga Arias, Mark Torrance, Thierry Olive, Rui A. Alves</i>	
The effect of task instruction and text overlap on the integration of multiple cross-domain sources... 73	
<i>Carla Firetto, Peggy Van Meter</i>	
Critical aspects of scientific phenomena – to the fore, in the background, or not present in scientific representations.....	76
<i>Tobias Fredlund, John Airey, Cedric Linder</i>	
A comparison of prompts, corrective feedback, and elaborative feedback in multimedia learning.....	79
<i>Hector Garcia Rodicio, Emilio Sanchez</i>	
What do representations say? – An analysis of students verbalizations.....	82
<i>Marion Geiger, Markus Vogel, Tina Seufert</i>	
Studying and executing procedures: Do pictures facilitate visuo-spatial reasoning during learning? ..	85
<i>Erlin van Genuchten, Charlotte van Hooijdonk, Katharina Scheiter, Anne Schüller</i>	
Picture-word learning by spacing the learning sessions .....	88
<i>Nicole Goossens, Gino Camp, Peter Verkoeijen, Huib Tabbers, Rolf Zwaan</i>	
What's the point of bulletpoints? The use of text in powerpoint lectures .....	91
<i>Madeline Hallewell</i>	
Learning with dynamic visualizations: the role of the human mirror neuron system .....	94
<i>Birgit Imhof, Ann-Christine Ehlis, Peter Gerjets</i>	
The impact of discrepancies across web pages on high-school students' trustworthiness evaluations .	97
<i>Yvonne Kammerer, Peter Gerjets</i>	
Does construction of diagrams deepen understanding by raising awareness of insufficiency in learning? .....	100
<i>Sachiko Kiyokawa, Yotan Kura, Yuri Uesaka, Emmanuel Manalo</i>	
The effects of visualization forms on usability and learning outcomes – dynamic videos versus static pictures .....	103
<i>Alexandra König, Mona Stadler, Melina Klepsch, Tina Seufert</i>	
Biology teachers' professional knowledge of diagrams.....	106
<i>Lena von Kotzebue, Claudia Nerdel</i>	
Learning art history on multi-touch-tables: metaphorical meaning of interaction gestures matters ...	109
<i>Julia Kranz, Birgit Imhof, Stefan Schwan, Barbara Kaup, Peter Gerjets</i>	
The misinterpretation of box plots .....	112
<i>Stephanie Lem, Patrick Onghena, Lieven Verschaffel, Wim Van Dooren</i>	
Secondary students' reading of digital visual representations when using physics educational simulations .....	115
<i>Victor López Simó, Roser Pintó Casulleras</i>	

Is making written syntheses an aid to the comprehension of documentary sources?.....	118
<i>Eva Lordán, Mariana Miras, Esther Nadal, Sandra Espino, Marta Minguela, Cristina Luna</i>	
Towards a suitable way to know what undergraduates believe about reading.....	121
<i>Eva Lordán, Francesc Salvador Beltrán, Isabel Solé</i>	
Using vibration to guide exploration of haptic graphics .....	124
<i>Richard K. Lowe</i>	
Addressing challenges of biological animations .....	127
<i>Richard K. Lowe, Jean-Michel Boucheix</i>	
Visualization and manipulation of English sounds on an interactive whiteboard at primary school..	130
<i>Emilie Magnat</i>	
What features make decorative illustrations interesting? .....	133
<i>Ulrike Magner, Rolf Schwonke, Inga Glogger, Alexander Renkl</i>	
Effects of text-belief consistency and reading goals on the comprehension of multiple science-related texts .....	136
<i>Johanna Maier, Tobias Richter</i>	
The role of epistemological beliefs and reading beliefs in multiple text comprehension .....	139
<i>Mar Mateos, Isabel Solé, Nuria Castells, Jara González Lamas</i>	
Effects of a mind map intervention on fifth and sixth graders' learning from texts .....	142
<i>Emmelen Merchie, Hilde Van Keer</i>	
Dynamic representations on the interactive whiteboard.....	145
<i>Jan van der Meij, Hans van der Meij</i>	
Does targeting the situation model during retrieval promote transfer from expository texts? .....	148
<i>Stijn van Mierlo, Huib Tabbers, Fred Paas</i>	
A simple scale of noun abstractness for predicting popular-scientific text comprehension.....	151
<i>Jaan Mikk</i>	
Meaningful rereading for meaning: the role of self-regulation in successful reading procedures .....	154
<i>Marta Minguela, Isabel Solé</i>	
Animations facilitate spatial perspective taking .....	157
<i>Stefan Münzer</i>	
Modeling the relationship between representational competence and domain knowledge.....	160
<i>Sandra Nitz, Claudia Nerdel, Helmut Precht</i>	
Measuring representational competence in science.....	163
<i>Sandra Nitz, Christine D. Tippet</i>	
Implicit and explicit guidance for collaborative learning at the museum .....	166
<i>Laure Philippon, Katharina Scheiter, Daniel Bodemer</i>	
Fourth graders' text and picture integration in processing and learning from science text: evidence from eye movement patterns .....	169
<i>Patrik Pluchino, Maria Caterina Tornatora, Lucia Mason</i>	
Effects of simplified ancillary representations and cues on learning from animation .....	172
<i>Dian-Kemala Putri, Jean-Michel Boucheix, Richard. K. Lowe</i>	
Visual ability in navigation communication.....	175
<i>Annemarie Quispel, Alfons Maes</i>	
Teenage students' awareness of source quality when selecting web links .....	178
<i>Jean-François Rouet, Christine Ros, Guillaume de Pereyra, Mônica Macedo-Rouet, Ladislao Salmeron</i>	
Selecting pages from Google to learn about a controversial topic: the role of epistemic beliefs.....	181
<i>Ladislao Salmerón, Yvonne Kammerer</i>	
Move like a fish: do gestures aid learning from photographs and videos? .....	184

<i>Katharina Scheiter, Jana Arndt, Birgit Imhof, Shaaron Ainsworth</i>	
Information access patterns of students working on text-picture integration tasks.....	187
<i>Wolfgang Schnotz, Ulrike Hochpoechler, Thorsten Rasch, Mark Ullrich, Holger Horz, Nele McElvany, Jürgen Baumert</i>	
Unwinding the relationship between cognitive processes and gaze behaviour during multimedia learning.....	190
<i>Carina Schubert, Katharina Scheiter, Anne Schüler</i>	
What pictures are good for .....	193
<i>Anne Schüler, Katharina Scheiter, Francesca Pazzaglia</i>	
Long-term multimedia learning.....	196
<i>Judith Schweppe, Ralf Rummer</i>	
Visual attention guidance in narrated animations: understanding is more than just seeing .....	199
<i>Irene T. Skuballa, Rolf Schwonke, Alexander Renkl</i>	
Does relevance matter in comprehending scientific conflicts from multiple documents? Evidence from online and offline-data.....	202
<i>Marc Stadler, Lisa Scharrer, Rainer Bromme</i>	
Do prior attitudes influence epistemic cognition while reading conflicting information? .....	205
<i>Johan L. H. van Strien, Saskia Brand-Gruwel, Henny P. A. Boshuizen</i>	
Not static or dynamic media but navigational affordances promote active learning .....	208
<i>Huib Tabbers, Jacob Diepenhorst</i>	
Benefits of a training for visualizing as a learning strategy .....	211
<i>Felix Wagner, David Rudolf, Tina Seufert</i>	
Metaphors activate object shape.....	214
<i>Lisanne van Weelden, Joost Schilperoord, Alfons Maes</i>	
Author Index .....	217



## KEYNOTE 1

### **The Early Development of Text Comprehension Skills: Emotional Inferences are Relevant Cues**

*Nathalie Blanc*

Epsilon Laboratory Dynamics of Human Abilities and Health Behaviors  
University of Montpellier 3  
Montpellier, France

Recent cognitive-developmental research shows that comprehension skills relevant to reading comprehension start developing early in children's lives - before children reach elementary school age. In this context, we examine how comprehension skills develop in preschool and early elementary school children. Because comprehension skills can be assessed through different media, we decided to use narrative-based televised stories to explore the development of one particular implicit comprehension skill: inference-making. This skill is essential to ensure good understanding of a text.

In a series of experiments, we studied children's ability to produce inferences from various depicted situations about characters' emotions and their potential causal links to story events. Taken together, our studies paint a picture indicating that emotional inferences are relevant cues to explore comprehension skills with preschool children.

The contribution of this series of experiments is threefold. First, at a theoretical level, the results of our studies are consistent with the recent cognitive-developmental model of comprehension skills (see van den Broek et al., 2005) according to which inference-making skills can be learned and assessed at a young age. Second, at a methodological level, we provide additional evidence that narrative-based televised stories offer relevant opportunities for early literacy-skill development. Third, we further support the idea that the skills acquired through repeated experience with narrative-based televised stories can also be transferred to understanding similarly constructed stimuli, like narrative-based storybooks. Overall, when used appropriately, narrative-based televised stories could be a useful tool for future research investigating activities that enhance young children's early literacy.

## KEYNOTE 2

# **Of Metaphors, Maps, and Methods: Cognitively Inspired and Perceptually Salient Graphic Displays**

*Sara Irina Fabrikant*

Geographic Information Visualization & Analysis (GIVA)  
Department of Geography, University of Zurich  
Zurich, Switzerland

In recent years, cartographers and GIScientists have become involved in extending geographic concepts and cartographic design approaches to the depiction of massive, non-geographic data archives. These so-called information spaces also incorporate explicit geographic metaphors with the intention to create a graphic representation that is easier to comprehend for information seekers.

In this presentation, I report on current progress and future opportunities in the emerging research field commonly known as spatialization or information visualization. More specifically, I propose an empirically validated design framework for the construction of cartographically sound spatialized displays based on spatial metaphors. As empirical studies on spatialized views suggest, basic geographic principles and cartographically informed design guidelines enable information designers to not only construct conceptually robust and usable information spaces, but also allow information seekers to more efficiently extract knowledge buried in large digital data archives.

I will highlight how knowledge from cognitive science and vision research can now help display designers systematically assess their designs, which are becoming increasingly dynamic and interactive. Methods such as eye tracking and the change blindness paradigm have been applied for that purpose, and cognitive/vision theories help make sense of the results, guiding the process of designing maps for salience and creating useful toolboxes. I will also discuss how cartographic methods themselves can help researchers make sense of data collected in user studies. This includes tools and methods such as sequence analysis and visual analytics.

## KEYNOTE 3

# **What Kind of Graphic is This? A Framework for Delineating the Graphics in Text-Graphic Research**

*Neil H. Schwartz*

Department of Psychology, College of Behavioral and Social Sciences  
California State University  
Chico (CA), USA

*Robert W. Danielson*

University of Southern California  
Los Angeles (CA), USA

Text-graphic research focuses on questions aimed at uncovering the way people understand to-be-learned material when text and graphics are presented together. The problem is that much of the research employs graphics whose properties often are not well delineated and frequently ill defined. Multiple taxonomies exist which attempt to categorize these properties, but the taxonomies are typically directed at some types of graphics and not others, and often neglect to take into account pertinent characteristics of learners (e.g. subject prior knowledge, and graphicacy skills), specification of learning task demands, and the requirements of learning outcomes.

In the present address, we will describe the importance of considering these dimensions in existing and future text graphic research, elucidating issues of redundancy, relativity, ambiguity, and context dependency of graphics relative to the text they accompany. Finally, we will present a framework for classifying and interpreting the text-graphic research by inventorying the findings of a number of published investigations, underscoring the necessity to account for these undifferentiated influences.

# JURE PRE-CONFERENCE WORKSHOP

## **An Eye Tracking Approach to Analyze Dynamic Representations**

*Jean-Michel Boucheix*

Laboratory for the Study of Learning and Development (*LEAD*)  
Burgundy University, Dijon, France

*Jonathan Groff*

French Institute of Science and Technology for Transport, Development and Networks (IFSTTAR)  
Lyon, France

In many studies, the effect of graphical representations on comprehension is evaluated by off-line tests (questionnaires etc...). These studies only measure the result of the learning process. So it's difficult to identify the particular elements which impact on learning processes. Eye tracking techniques make it possible to give on-line information about learners' behaviour during the experiment. In SIG 2, graphical representations are studied in terms of their cognitive effects. So the aim of this workshop is to give an insight into the possibilities and limitations of eye tracking within the areas of learning and comprehension of graphical representations. In this workshop, we propose to demonstrate how eye tracking analysis can reveal cognitive processes involved in the research of information, in the extraction of relevant information and in the elaboration of a dynamic mental model. Three examples will be used to illustrate this analysis technique: the upright piano mechanism, animated public information graphics and fish locomotion. During this session, we will present the contribution of eye tracking in the studies of representation processes. Then, participants will learn to use an eye tracker (Tobii t120) with the three examples previously mentioned.

# Effects of Constructing Concept Maps while Navigating in a Hypertext<sup>1</sup>

*Franck Amadiou, Julie Lemarié, Aline Chevalier*

CLLE-LTC, University of Toulouse, 5, Allées A. Machado, 31058 Toulouse Cedex, France  
amadiou@univ-tlse2.fr, lemarie@univ-tlse2.fr, aline.chevalier@univ-tlse2.fr

*Julien Cegarra*

CLLE-LTC, Jean-François Champollion University Centre  
Centre Universitaire Jean-François Champollion, Place de Verdun, 81012 ALBI Cedex 09, France  
Julien.Cegarra@univ-jfc.fr

*Ladislao Salmerón*

ERI Lectura & Department of Developmental and Educational Psychology, University of Valencia  
Avda. Blasco Ibáñez, 21, 46010 Valencia, Spain  
Ladislao.Salmeron@uv.es

*Stéphane Blat*

High school Toulouse Lautrec, 64 Boulevard Pierre et Marie Curie 31200 Toulouse, France  
stephane-christ.blat@ac-toulouse.fr

**Abstract.** The present study aims to investigate the effects of prior domain knowledge and of constructing a concepts map on hypertext comprehension. Three conditions of a hypertext dealing with the greenhouse effect were compared: (a) a linear map (coherent order of concepts), (b) a system-structured map, and (c) a self-structured map (participants had to construct themselves the map). The results failed to show that a self-structure concept map would promote active learning and deep processing for high prior knowledge readers. However, it was confirmed that low prior knowledge readers benefited from the system-structured concept map that stresses the hypertext macrostructure.

**Keywords:** Active learning, Concepts map, Cognitive load, Hypertext reading, Prior domain knowledge

## Introduction

Within the research field on learning from hypertext, concept maps have been the topic of a pronounced interest. A concept map is a graphic organiser displaying concepts and links (labelled or not and directional or not) that show semantic relations between concepts. Concept maps supply guidance for hypertext comprehension (e.g. Waniek, 2012). Actually, processing a concept map would support comprehension especially when hypertext contents are difficult (Salmerón, Baccino, Cañas, Madrid, & Fajardo, 2009).

According to the active processing model (Hofman & van Oostendorp, 1999; Salmerón et al., 2009), concept maps conveying the hypertext macrostructure may also inhibit deep comprehension strategies for high prior knowledge (HPK) readers. When the concepts map explicitly conveys the text structural information, HPK don't need to be active to create the macrostructure of the text information. Therefore, the aim of the present study is to investigate the effect of a new type of concept map that would foster deep processing of hypertexts by HPK readers.

In the research field of learning from paper documents, according to Stull and Mayer (2007), constructing concept maps may increase the level of extraneous cognitive processing (i.e. inefficient processes for learning) and limits generative processing as regard to the Cognitive Load Theory; conversely, challenging readers to construct their own concept map structure may promote active reading as regard to the activity theory (Klahr & Nigam, 2004). However, in the field of learning with hypertexts, a research by Amadiou, Tricot and Mariné (2009) showed that learners' prior knowledge was linked to higher comprehension performance, more coherent reading orders, more efficient processes to construct the document structure and lower experienced cognitive load.

In the present study, structuring a concept map during a comprehension task is expected to promote relational and structural processes for high prior knowledge (HPK) readers. Their reading of hypertext

---

<sup>1</sup> Project supported by ANR (National Research Agency)

would be more active through deeper processing of both contents and semantic organization. Conversely, a self-structured concept map would entail a dual task for low prior knowledge (LPK) readers. It will impose additional processing that will increase extraneous cognitive load and consequently hamper comprehension. A system-structured concept map should support their comprehension in a greater extent than a self-structured concept map.

## **Method (experiment in progress and finished at the end of April 2012)**

### *Participants.*

263 French high school students (10<sup>th</sup> grade) took part to the experiment (mean age = 15.48).

### *Materials*

*Learning task and materials.* A hypertext dealing with the greenhouse effect was designed. The learning material consisted in twelve concepts (12 pages with a total number of words = 596). Three concept maps have been designed: (a) a linear map (a coherent structure that respected the sequence of the greenhouse mechanism), (b) a system-structured map (a map displaying the main relations between concepts that could not be modified by participants), and (c) a self-structured concept map (the concepts were provided and the participants had to structure the map moving the concepts and drawing labelled links between them).

Concept labels and hypertext sections were identical for the three map conditions. In all the concept map conditions, clicking on a concept opened a text section dealing with this concept. Then a link below the text brought back to the map where a new concept (or the same concept) could be opened. Participants were free to consult the text sections several times and in a free order.

*Prior knowledge tests.* Prior general domain knowledge was assessed by 9 multiple choice questions about general physic principles that were useful to learn about the greenhouse effect. Prior subject-matter knowledge (pre-test) was assessed with the same questions than the one used in the comprehension test. From the pre- and post-test, a gain of knowledge was measured (post-test – pretest). For details about the pre and post-test questions, see the comprehension performance below.

*Comprehension performance.* 12 questions required inferences between different text sections (inference questions) while 12 other questions dealt with information explicitly mentioned in text sections (textbased questions).

*Cognitive load and disorientation measures.* For each major part of the learning task, participants had to rate their mental effort, perceived difficulty to understand the text section and perceived complexity to understand the relations between the text sections (7- point Likert scale from very low to very high).

*Controlled variables.* Three variables were assessed for control: (a) reading skills (decoding test), (b) individual interest for the topic and (c) scanning reading strategies of web documents for studying.

### *Procedure*

The participants filled in questionnaires and rated scales in the following order: reading skills, personal interest, prior domain knowledge, reading strategies habits. In the same session, they performed next a training task with the interactivity functions with the document. The learning task was limited to 20 minutes. Finally, participants rated their mental effort and their perceived complexity, and performed the comprehension test.

## **Results**

A 2x3 (level of prior knowledge X type of map) ANCOVA with three covariates (individual interest, reading skills, scanning reading strategies on internet) were conducted to analyse the results.

### *Comprehension scores*

The analyses conducted on the textbased questions scores indicated that the HPK readers outperformed the LPK readers;  $F(1, 253) = 15,11$ ,  $p < .001$ , and positive effects of the 3 covariates: interest ( $p = .004$ ), reading skills ( $p = .013$ ), scanning strategy ( $p = .001$ ).

The analyses on the inferential questions scores revealed also a main effect of prior knowledge;  $F(1, 251) = 22.34$ ,  $p < .001$ , a marginal main effect of the map,  $F(2, 251) = 2.70$ ,  $p = .069$  and a marginal interaction effect,  $F(2, 251) = 2.63$ ,  $p = .074$ . Pairwise comparisons revealed an effect of the map only for the prior knowledge readers,  $F(2, 251) = 5.20$ ,  $p = .006$ , showing that the system-structured map supported better scores than the linear ( $p = .071$ ) and the self-structured map ( $p = .001$ ). The interest ( $p = .005$ ) as well as the scanning strategy ( $p = .001$ ) positively impacted the inferences questions scores.

Table 1: Comprehension scores, mental effort and perceived complexity scores.

			Textbased questions scores	Inferential questions scores	Mental effort ratings	Complexity: text sections	Complexity: relationships
Linear map	LPK	Mean(SD)	6,45 (3,05)	6,45 (3,05)	4,32 (1,19)	3,63 (1,66)	4,20 (1,54)
	HPK	Mean(SD)	7,75 (2,48)	7,75 (2,48)	3,92(1,13)	3,29 (1,11)	3,63 (1,35)
Self-structured map	LPK	Mean (SD)	6,04 (2,40)	6,04 (2,40)	4,27 (1,36)	3,65 (1,21)	4,17 (1,42)
	HPK	Mean (SD)	7,70 (2,96)	7,70 (2,96)	4,13 (1,02)	3,24 (1,48)	3,98 (1,11)
System- structured map	LPK	Mean (SD)	6,49 (2,35)	6,49 (2,35)	4,51 (1,08)	3,71 (1,23)	3,90 (1,22)
	HPK	Mean (SD)	7,47 (2,60)	7,47 (2,60)	4,05 (,99)	3,22 (1,25)	3,00 (1,13)

### *Mental effort and perceived complexity*

The analyses only showed that LPK experienced higher mental effort to learn the mechanism of the greenhouse effect,  $F(1,254) = 5.98$ ,  $p = .015$ . An similar results pattern was observed for the perceived complexity related to the comprehension of the text sections (main effect of prior knowledge,  $F(1,254) = 5.69$ ,  $p = .018$ ). The scanning strategy had also a positive significant effect on reported mental effort,  $F(1,254) = 5.93$ ,  $p = .016$ . However, no others effects were observed on the perceived complexity to understand the relationships between the concepts. A main effect of prior knowledge is still significant,  $F(1,254) = 12.32$ ,  $p = .001$ . A main effect of the type of map indicated that the system-structured map reduced the complexity to understand the relationships in comparison to the linear structure ( $p = .039$ ) and to the self-structured map ( $p = .001$ ). Interest ( $p = .025$ ) and scanning reading strategy ( $p < .001$ ) had a positive effect also.

### **Conclusion**

The results did not reveal a positive effect of the self-structure map for HPK learners. Nevertheless, the analyses on the comprehension performance indicated that the HPK readers were not affected by the type of map. The results of the LPK readers were consistent with our expectations. The system-structured map supported a higher comprehension of the relations between the concepts (macrostructure) than the linear and the self-structured map. The ratings on the measures of cognitive load showed that the system-structured map limited the cognitive requirements to understand the implicit relations between the concepts for all the participants. Analyses of on-line data will be presented in the poster.

### **Bibliography**

- Amadiou, F., Tricot, A., & Mariné, C. (2009). Exploratory Study of Relations Between Prior Knowledge, Comprehension, Disorientation and On-Line Processes in Hypertext. *The Ergonomics Open Journal*, 9, 49-57.
- Hofman, R., & Oostendorp, H. V. (1999). Cognitive effects of a structural overview in a hypertext. *British Journal of Educational*, 30(2), 129-140.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, 15, 661-667.
- Waniek, J. (2012). How information organisation affects users' representation of hypertext structure and content. *Behaviour & Information Technology*, 31(2), 143-154.
- Salmerón, L., Baccino, T., Cañas, J. J., Madrid, R. I., & Fajardo, I. (2009). Do graphical overviews facilitate or hinder comprehension in hypertext? *Computers & Education*, 53(4), 1308-1319.
- Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology*, 99(4), 808-820.

# Visualizing Football Statistics: Performance & Preference

*Marije van Amelsvoort*

*Hans Westerbeek*

Tilburg center for Communication and Cognition, Tilburg University,  
Warandelaan 2, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands  
m.a.a.vanamelsvoort@uvt.nl  
h.g.w.westerbeek@uvt.nl

**Abstract.** Two visualizations of the statistics of a football match are compared that differ in terms of reality, interactivity, and spatial organization. Overall performance on reading off information was the same in both visualizations, but simple questions were more often answered correctly in the conventional Bar view, while questions that needed combining information were more often answered correctly in the more realistic and interactive Pitch view. Age interacted with performance on the two views as well. Preference for one of the two visualizations depended on the type of task; while the Bar view was considered more suitable to remember information or get an overview, the Pitch view was preferred to watch the information for fun or see the course of the game. Unconventional designs may be more fun, but they are only more effective for certain types of tasks.

**Keywords:** external representations, information design, performance, preference

## Introduction

During the 2010 FIFA Football World Cup, the BBC presented two types of visualizations of match statistics on their website. These visualizations were informationally equivalent (Larkin & Simon, 1987), but differed in terms of spatial organization and interactivity. The Bar view - a more conventional and domain independent representation of football statistics - showed a list of players, cards, and substitutions plus horizontal bars with proportions of possession, corners and the like. The Pitch view showed the same information, but made use of a spatial layout that mimicked a football pitch with a clock in the middle on which game events were set out. In this view, some information only appeared when hovering over the pitch with the mouse pointer (see Figure 1 for both views). This presentation of information led us to two questions: which view is better for finding information (performance), and what view do people prefer (preference)?

The Pitch view seems to have several advantages over the Bar view for performance in searching for information about game statistics. First, there is a more natural mapping of information on the football field (albeit seen from above). For example, information about the amount of corners taken are placed in the corner, which should facilitate search processes (Hegarty, 2011). Second, the clock gives an overview of the course of the game, which should facilitate encoding information (Baddeley, Eysenk, & Anderson, 2009). Third, the interactivity gives users a sense of control, which should lead to more cognitive processing (Mayer & Moreno, 2003).

However, there are also reasons to expect that the Bar view is more effective for understanding football matches. The Bar view is more conventional, which may make it easier to work with, especially for older people not used to highly interactive graphics, or for people who often look at match statistics. Also, basic knowledge about football is needed to reap the advantages of the perceptual properties of the Pitch view (Kosslyn, 2006). In addition, performance with a more realistic representation is often not so good, even though people have a strong preference for realistic displays (Smallman & St. John, 2005).

Not only user characteristics, but also type of task may influence performance and preference for either one of the views. Since the Pitch view is interactive and shows a clear timeline of events, we expect that this view may be helpful when searching for combined information about the match, e.g., combining event data (what happened) with time data (when did it happen). In addition, we hypothesize that the Pitch view may be preferred over the Bar view for tasks dealing with fun (due to



layout and interactivity), or with understanding the course of the game (due to the clock), while the Bar view may be preferred for tasks that involve quickly reading off information.

## Method

541 participants took part in this study. Their mean age was 30.9 years ( $SD=14.6$  years). Two age groups could be distinguished, one group of 40 years and younger ( $N=375$ ) and one group of 41 years and over ( $N=166$ ). Participants first filled a questionnaire on demographics, relevant experience with football, match statistics and the internet. Then they viewed two matches each (one in the Pitch view and one in the Bar view). They answered fourteen questions about the match's statistics. Nine of these questions inquired simple information readily available in both view types (e.g.: "what was the final score of the match?"). The answers to these questions could be directly read from the displays. The other five questions were so-called complex questions, and required different pieces of information to be combined (e.g.: "who scored the winning goal in this match?"). We measured correctness and response times. After this questions task, participants compared the two views on clarity, usefulness and completeness. The same choice had to be made for seven usage scenarios, inquiring about which view people preferred when they, for example, had to remember the match information or when they wanted to see the development of the match.

## Results

### *Correctness*

There was no main effect of view type on the percentage of correctly answered questions,  $F<1$ , Bar = 77%; Pitch = 78%. However, view type interacted with question type,  $F(1,540)=38.03$ ,  $p<.001$ ,  $\eta^2=.066$ . Using the Pitch view, participants answered significantly more simple questions correctly, 79%, than complex questions, 77%,  $F(1,540)=6.98$ ,  $p<.01$ ,  $\eta^2=.013$ . Using the Bar view, the same effect of question type was found, but it was significantly stronger (81% vs. 72%),  $F(1,540)=115.74$ ,  $p<.001$ ,  $\eta^2=.177$ . For simple questions the Pitch view induced fewer correct answers than the Bar view, 79% vs. 81%,  $F(1,540)=5.18$ ,  $p<.025$ ;  $\eta^2=.009$ , and for complex questions the situation was vice versa,  $F(1,540)=20.17$ ,  $p<.001$ ,  $\eta^2=.036$ . View type also interacted with age group,  $F(1,538)=6.56$ ;  $p<.025$ ;  $\eta^2=.012$ . The younger participants outperformed their older peers using the Pitch view, but there were no differences between the age groups for the Bar view.

### *Response Time*

Participants responded significantly slower when using the Pitch view than when using the Bar view,  $F(1,538)=227.22$ ;  $p<.001$ ;  $\eta^2=.297$ , both for simple and for complex questions. View type interacted with question type,  $F(1,538)=5.89$ ;  $p<.025$ ;  $\eta^2=.011$ . In both views, participants responded significantly quicker on simple questions than on complex questions. The younger participants responded significantly quicker than the older participants in both views.

### *Preference*

The Bar view was found clearer by 69.4% of the participants,  $\chi^2(1)=80.70$ ,  $p<.001$ , more useful, 69.4%,  $\chi^2(1)=80.07$ ,  $p<.001$ , and more complete, 54.9%;  $\chi^2(1)=4.95$ ,  $p<.05$ , than the Pitch view. The choices participants made in the seven usage scenarios indicate that the Bar view is preferred when they had to remember the information, 60.5%;  $\chi^2(1)=23.30$ ,  $p<.001$ , to get an overview of the match, 68.8%;  $\chi^2(1)=74.53$ ,  $p<.001$ , and to quickly view the match, 67.9%;  $\chi^2(1)=67.12$ ,  $p<.001$ . The Pitch view was judged more suitable for watching the information for fun, 69.3%;  $\chi^2(1)=78.20$ ,  $p<.001$  and to see the course of the match, 66.3%;  $\chi^2(1)=56.03$ ,  $p<.001$ . No clear preference for one of the views was expressed for explaining the match to someone else and for wanting to understand the match.

## Discussion

Designers often aim for visualizations that are attractive and new, which may come at a cost of being unfamiliar and difficult for users. The unconventional and interactive Pitch design in this study was hard for everyone, especially for older people, considering the differences in response time. A representation with a more natural mapping is not automatically better, especially in this example where it is combined with abstract information that is represented symbolically (e.g., the size of the circle). Users took type of task into account in deciding which view was better. Unconventional designs seem to be regarded as more fun, but not necessarily as more effective.

The Pitch view differs from the Bar view in two essential ways: it is both more spatial and interactive. To distinguish between those two variables, we conducted a follow-up experiment in which we created two Pitch views: one with and one without interactivity. We also conducted a production experiment in which participants were asked to write a report of the match to a friend who wasn't able to see the game. We hypothesize that the two views would lead to different descriptions in terms of story line. Results of these experiments will be discussed in Grenoble.

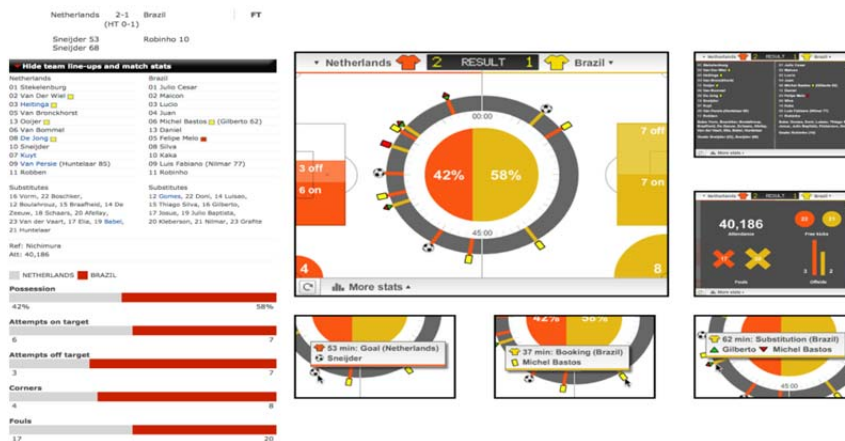


Figure 1. Bar view (left) and Pitch view (right). The figures on the right are shown when hovering over Pitch, the figures below when hovering over the clock.

## References

- Baddeley, A., Eysenck, M.W., & Anderson, M.C. (2009). *Memory*. London: Psychology Press, pp. 27, 44-59.
- Hegarty, M. (2011). The cognitive Science of Visual-Spatial Displays: Implications for design. *Topics in Cognitive Science*, 3, 446-474.
- Kosslyn, S. M. (2006). *Graph design for the eye and mind*. New York: Oxford University Pres.
- Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Mayer, R. E., & Moreno, (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43-52.
- Smallman, H. S., & St. John, M. (2005). Naïve Realism: Misplaced faith in realistic displays. *Ergonomics in design*, 13, 14-1.

# Metacognition in Reading Comprehension: Descriptive Study

*Olga Arias, Raquel Fidalgo*

Universidad de León. Departamento de Psicología, Sociología y Filosofía  
Campus de Vegazana s/n, 24071 León, España  
o.arias.gundin@unileon.es, rfidr@unileon.es

**Abstract.** Reading comprehension is a complex task that depends on many different automatic and strategic cognitive processes. The aim of this study is to explore differences in metacognition in reading comprehension. Participants were elementary school students from fifth to sixth grade in primary education and from first to second grade in secondary education. A metacognitive knowledge questionnaire was used to measure metacognition in reading comprehension. The results obtained in this study showed the improvements of metacognitive knowledge developed during elementary school; the explicit reference to decoding knowledge decreases at the same time semantic knowledge increases.

**Keywords.** Reading, comprehension, metacognition, primary education, secondary education

## Background

Research within the domain of reading comprehension has led to an increasing emphasis on the role of metacognitive awareness of one's cognitive and motivational processes while reading; specifically researchers agree that awareness and monitoring of one's comprehension processes are critically important aspects of skilled reading (Sen, 2009). The central role of metacognition and comprehension monitoring in the current descriptions of the reading process is reflected in the steady growth of interest in reading comprehension monitoring research.

Comprehension monitoring is important for the regulation of reading that is manifested in the way readers plan, monitor, evaluate, and use information available to them as they make sense of what they read. Changes in comprehension monitoring and in the use of reading strategies are aspects of metacognitive development which begins with young children's awareness of mental functions, and eventuates in complex metacognitive abilities which many adults do not master (Kuhn, 2000). During development, metacognition becomes more explicit, powerful, and effective, and it comes to operate increasingly under the individual's conscious control (Eme, Puustinen, & Coutelet, 2006). Several studies indicate that elementary school is a critical period of the development of metacognition and comprehension monitoring flexibility (Kolic-Vehovec & Bajanski, 2006).

## Aim

The aim of this study is to examine differences in self-knowledge reading comprehension in students from 5<sup>th</sup> Primary Education to 2<sup>nd</sup> Secondary Education.

## Method

### *Participants*

The participants in this study were students (N=509) from fifth to sixth grade of primary education (aged 10-12) and from first to second grade of secondary education (aged 12-14) in five elementary schools in León, Spain: 117 fifth-graders (59 girls and 58 boys), 134 sixth-graders (66 girls and 68 boys), 123 first-graders (59 girls and 64 boys) and 135 second-graders (70 girls and 65 boys).

### *Measures*

A metacognitive knowledge was measured with a questionnaire, which consists of six open-questions. In order to analyse and categorize all answers, a system of comprehensive and mutually exclusive

categories was developed following a rational and empirical system. We took into account as central axis, the difference between reading levels, strategic knowledge, self-regulation and other moderator variables of reading comprehension (e.g. motivation, previous knowledge, purpose...).

### Procedure

The questionnaire was distributed to the students in their classrooms as whole groups after they had read and summarized a text.

### Results

Means and standard deviations for all grades on each categories of metacognitive knowledge assessment are presented in Table 1.

Table 1: Means and standard deviations for all grades on each categories of metacognitive knowledge

	5 <sup>o</sup> (N=117)		6 <sup>o</sup> (N=134)		1 <sup>o</sup> (N=123)		2 <sup>o</sup> (N=135)	
Categories	M	SD	M	SD	M	SD	M	SD
Decoding level	.1238	.1156	.0904	.0962	.0762	.0899	.0759	.0880
Syntactic level	.0288	.0510	.0208	.0402	.0302	.0532	.0251	.0413
Semantic level	.0982	.0411	.1022	.0473	.1137	.0404	.1081	.0397
Situational level	.0232	.0469	.0263	.0455	.0201	.0432	.0218	.0396
Strategic knowledge	.1436	.0979	.1514	.0962	.1424	.0916	.1481	.0871
Self-regulation	.0773	.0883	.0793	.0860	.0938	.0895	.0955	.0906
Motivation	.0557	.0731	.0712	.0777	.0611	.0679	.0787	.0745
Previous knowledge	.0048	.0175	.0019	.0094	.0041	.0159	.0056	.0182
Reader's purpose	.1086	.0699	.1106	.0623	.1086	.0522	.1044	.0536
No reader's purpose	.0033	.0128	.0038	.0164	.0000	.0000	.0002	.0029
Efficacy	.0026	.0119	.0023	.0094	.0041	.0124	.0034	.0124
Practice	.0695	.0814	.0813	.0829	.0732	.0796	.0705	.0728
Self-efficacy	.0026	.0119	.0023	.0094	.0041	.0124	.0034	.0124
Ability	.1078	.0780	.1043	.0734	.1118	.0786	.1131	.0815
Assessment text	.0150	.0449	.0123	.0346	.0186	.0366	.0142	.0299
Assessment context	.0043	.0222	.0073	.0255	.0059	.0254	.0027	.0105
Uncategorizable	.0301	.0469	.0263	.0446	.0179	.0365	.0226	.0417

Differences in a metacognitive knowledge between grades were tested by MANOVAs  $\lambda = .861$   $F(3,507) = 1.564$ ,  $p < .009$ . We found statistically significant differences in decoding level  $F(3,507) = 6.433$ ,  $p < .001$ , semantic level  $F(3,507) = 3.155$ ,  $p < .025$ , and no reader's purpose  $F(3,507) = 4.650$ ,  $p < .003$ .

In post-hoc comparisons we observe in decoding level differences between the students of the 5<sup>th</sup> grade ( $M_{5th} = .1238$ ) with regard to the rest of the grades  $M_{6th} = .0904$ ,  $p < .007$ ,  $M_{1st} = .0762$ ,  $p < .001$ ,  $M_{2nd} = .0759$ ,  $p < .001$ . In semantic level we only observe differences between the 1<sup>st</sup> grade ( $M_{1st} = .1137$ ) with regard the 5<sup>th</sup> grade  $M_{5th} = .0982$ ,  $p < .005$  and the 6<sup>th</sup>  $M_{6th} = .1022$ ,  $p < .029$ . In no reader's purpose we observe differences in the 5<sup>th</sup> grade ( $M_{5th} = .0033$ ) with regard the 1<sup>st</sup> grade  $M_{1st} = .0000$ ,  $p < .013$  and the 2<sup>nd</sup> grade  $M_{2nd} = .0002$ ,  $p < .019$ , and in the 6<sup>th</sup> grade ( $M_{6th} = .0038$ ) with regard the 1<sup>st</sup> grade  $M_{1st} = .0000$ ,  $p < .004$  and the 2<sup>nd</sup> grade  $M_{2nd} = .0002$ ,  $p < .006$ .

Although no statistically significant differences were found in others categories, we want to highlight several aspects: firstly, the knowledge and importance assigned to self-regulation is increased with the grade; secondly, the high scores obtained, in each grade, both strategic knowledge

and setting the goal of reading; and thirdly the lack of importance that every students assigned to prior knowledge.

## **Conclusions**

Besides automatized basic reading processes, skilled reading also requires the ongoing monitoring comprehension, and regulation according to the goals of reading accomplished by the use of reading strategies (Alexander & Jetton, 2000).

The results obtained in this study show improvements in the development of metacognitive knowledge during elementary school. The results indicate that significant transition happens before the first grade, but there is not more transition periods since the sixth grade in different aspects of metacognition knowledge because there are no differences between students of the first and the second grade. The developmental effect on decoding level is inverse to semantic level, that is, the explicit reference to decoding knowledge decreases at the same time semantic knowledge increases. The same trend is observed in report to the lack of reader's purpose, it decreases until it disappears in the first grade.

This research was possible with funds from Science and Innovation Ministry of Spain through research project (EDU2010-18219), that was awarded to Dra. Fidalgo.

## **References**

- Alexander, P.A., & Jetton, T.L. (2000). Learning from text: A multidimensional and developmental perspective. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (vol. 3, pp. 285-310). Mahwah, NJ: LEA.
- Eme, E., Puustinen, M., & Coutelet, B. (2006). Individual and developmental differences in reading monitoring: When and how do children evaluate their comprehension? *European Journal of Psychology of Education*, 21 (1), 91-115.
- Kolic-Vehovec, S., & Bajanski, I. (2006). Metacognitive strategies and reading comprehension in elementary-school students. *European Journal of Psychology of Education*, 21 (4), 439-451.
- Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, 9, 178-181.
- Sen, S.H. (2009). The relationship between the use of metacognitive strategies and reading comprehension. *Procedia Social and Behavioral Sciences*, 1, 2301-2305.

# Examining the Integration of Text and Pictures

Jana Arndt, Anne Schüller, Katharina Scheiter

Knowledge Media Research Center

Schleichstraße 6, 72076 Tübingen, Germany

{j.arndt, a.schueler, k.scheiter} @iwm-kmrc.de

**Abstract.** According to the Cognitive Theory of Multimedia Learning (Mayer, 2009) one crucial step for learning with multimedia is the integration of text and pictures. The aim of the reported studies was to investigate this integration process by using a modified paradigm introduced by Gentner and Loftus (1979). Learners memorized series of pictures and sentences that resulted from cross-varying type of picture (general vs. specific) and type of sentence (general vs. specific) within subjects. It was assumed that if text and pictures are integrated with each other, learners should be less able to trace back the learned information to a specific representation. This should result in lower performance in a later recognition test. The results of the reported studies support the assumption of an integration process. However, it remains still unclear, whether the integrated model is verbal, pictorial, or has no specific codality.

**Keywords:** multimedia, integration process, text and pictures

## Introduction

According to the Cognitive Theory of Multimedia Learning (CTML, Mayer, 2009) the most crucial step for learning with text and pictures is the integration of verbal and pictorial information with each other and with prior knowledge. However, so far, there have been no studies explicitly investigating this integration process. Thus, the aim of the three studies reported below was to investigate the integration of text and pictures. For that purpose, we used a paradigm introduced by Gentner and Loftus (1979). Modified versions of this paradigm were also used by Pezdek (1977) and in research about eyewitness testimony (e.g., Belli, 1988; Loftus, 1975, 1977).

The basic idea of the paradigm is to present series of general and specific pictures, matched with series of general and specific sentences. For instance, subjects see a picture of a map. The depicted map includes either a treasure (specific picture) or not (general picture, see Figure 1). The two pictures are matched with the sentence “The *map* is really inaccurate” (general sentence) or the sentence “The *treasure map* is really inaccurate” (specific sentence). Thus, the specific sentence and the specific picture contain additional information compared to the general sentence and the general picture. If integration takes place, the specific information and the general information should be integrated with each other. As a consequence, learners should falsely recognize the specific information in a later recognition test. For instance, learners see the general picture without the treasure accompanied by the specific sentence about the treasure map. We assume that the learners integrate the general picture information about the map with the specific sentence information about the treasure map. Due to this integration process, learners should falsely recognize the specific picture containing a treasure and falsely reject the general picture without the treasure in a later recognition test.

Thus, we assume that the presentation of general sentences together with specific pictures (and vice versa) leads to worse performance in a recognition test compared to the presentation of general pictures together with general sentences or specific pictures together with specific sentences.

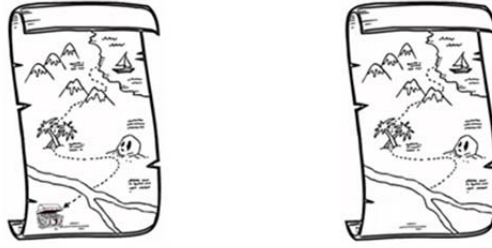


Figure 1. Specific picture (left) shows a treasure map; general picture (right) shows a map.

## Method and Results

This assumption was tested in three experiments. Learners were always instructed to memorize pairs of pictures and sentences that resulted from cross-varying type of picture (general vs. specific) with type of sentence (general vs. specific) within subjects. As dependent variable performance in a forced-choice recognition test was measured. In this test, pairs of pictures (one specific, one general) or pairs of sentences (one specific, one general) were presented to learners and they had to indicate which of the pictures or sentences had been presented in the learning phase. The three experiments differed concerning the number of presented stimuli and stimuli presentation time. To test our assumption,  $2 \times 2$  ANOVAs with the within-subject factors sentence type (general vs. specific) and picture type (general vs. specific) were conducted. Due to space limitations, we do not report main effects in the following, but focus on the interaction, which indicates whether text-picture integration takes place. That is, if integration took place, then students should show worse recognition performance in the conditions, where specific sentences are combined with general pictures and vice versa, compared with conditions where both representations have the same level of specificity.

In Study 1, 47 learners memorized 32 pairs of pictures and sentences. Every picture-sentence-pair was presented for 8 seconds. With regard to sentence recognition there was no interaction. With regard to picture recognition a marginal significant interaction appeared,  $F(1, 46) = 2.91$ ,  $MSE = 519.65$ ,  $p = .095$ ,  $\eta_p^2 = .06$ . Bonferroni post-hoc comparisons indicated that combinations of specific sentences and general pictures ( $M = 72.34$ ,  $SE = 4.74$ ) led to worse recognition of pictures than combinations of specific sentences and specific pictures ( $M = 86.53$ ,  $SE = 3.68$ ;  $p = .018$ ). This finding corroborates our assumption of an integration process: Specific sentence information was integrated with the general picture information, leading to an underlying mental model, which included the specific information. This in turn led to falsely recognizing the specific picture in the posttest.

The aim of Study 2 was to replicate the observed integration process by using a higher number of items. Therefore, in the second study, 75 learners memorized 59 pairs of pictures and sentences. Again, every picture-sentence-pair was presented for eight seconds. Contrary to Experiment 1, with regard to picture recognition no interaction was observed, whereas with regard to sentence recognition a significant interaction appeared,  $F(1, 74) = 5.18$ ,  $MSE = 294.20$ ,  $p = .026$ ,  $\eta_p^2 = .07$ . Bonferroni post-hoc comparisons indicated that combinations of general sentences and specific pictures ( $M = 82.62$ ,  $SE = 2.42$ ) led to worse recognition of sentences than combinations of general sentences and general pictures ( $M = 88.44$ ,  $SE = 1.75$ ,  $p = .033$ ). These results indicate that the specific picture information was integrated with the general sentence information, leading to an integrated model containing the specific picture information. However, unexpectedly, Bonferroni post-hoc comparisons also indicated that combinations of specific sentences and general pictures ( $M = 73.98$ ,  $SE = 2.47$ ) led to worse recognition of sentences than combinations of general sentences and general pictures ( $p < .001$ ). Thus, it seems that the integration process is not a unidirectional process, but that integration might also occur the other way around (i.e., general information can overwrite the specific information).

The aim of Study 3 was to analyse whether the observed integration process might be more pronounced with shorter presentation time. Therefore, every picture-sentence pair was presented for 5

seconds (instead of 8 seconds as in Study 1 and 2). Additionally, the number of items was increased once again. Forty-seven learners memorized 72 pairs of pictures and sentences. With regard to sentence recognition a marginal significant interaction was observed,  $F(1, 46) = 2.89$ ,  $MSE = 286.48$ ,  $p = .096$ ,  $\eta_p^2 = .06$ . Bonferroni post-hoc comparisons indicated that combinations of general sentences and specific pictures ( $M = 73.90$ ,  $SE = 2.56$ ) lead to worse recognition of sentences than combinations of specific sentences and specific pictures ( $M = 83.52$ ,  $SE = 1.98$ ;  $p = .002$ ). Moreover, combinations of specific sentences and general pictures ( $M = 78.00$ ,  $SE = 2.29$ ) lead marginally to worse recognition of sentences than combinations of specific sentences and specific pictures ( $p = .065$ ). Regarding picture recognition no interaction was observed. To sum up, in line with the results of Experiment 2, the integration effect was only observed for sentence recognition. Moreover, the finding of a bidirectional integration process was replicated, because general pictures presented together with specific sentences lead to worse sentence recognition, indicating that the general pictures information was incorporated into the mental model.

## Discussion

To summarize, the results support the assumption of CTML that text and picture are integrated with each other. However, it has to be noted that the manifestation of the integration process differed between experiments. Whereas in Experiment 1 the predicted pattern was observed concerning pictorial recognition, in Experiment 2 & 3 it emerged concerning sentence recognition. One explanation might be that the resulting mental model itself is abstract, leading to worse picture as well as sentence recognition in a post-test. Another explanation might be that in Experiment 2 & 3 participants stopped integration after some time and only looked at the pictures due to the higher item number and the shorter presentation times. This question should be answered in further research (e.g., by collecting process data). Furthermore, in Experiment 2 and 3 not only the specific picture information was integrated with the general sentence information but also the general picture information was integrated with the specific sentence information. This indicates that the integration process might be bidirectional. Further research is needed to clarify this question.

## References

- Belli, R. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses? *Memory & Cognition*, 16, 314-326.
- Gentner, D., & Loftus, E. F. (1979). Integration of verbal and visual information as evidence by distortions in picture memory. *The American Journal of Psychology*, 92, 363-375.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 550-572.
- Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition*, 5, 696-699.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed). New York: Cambridge University Press.
- Pezdek, K. (1977). Cross-modality semantic integration of sentence and picture memory. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 515-524.
- Pezdek, K. (1980). Life-span differences in semantic integration of pictures and sentences in memory. *Child Development*, 51, 720-729.



# Categorization of Graphical Representations

*Muhammad Ashraf*

Laboratory of Educational Sciences  
Université Pierre-Mendès-France, Grenoble, France  
Muhammad.Ashraf@upmf-grenoble.fr

**Abstract.** Role of graphics is significant irrespective of differences in domains. To great extent, information is presented through diagrams, pictograms, maps and charts on paper and digitally. We selected 50 graphics from different newspapers, textbooks, with information about geography, physics, chemistry and mathematics etc. The set of 50 graphics were categorized by two raters and Cohen's kappa test was computed to check the reliability. Subsequently, a card sorting task was conducted with 20 Master students using these graphics. A Cluster analysis showed that maps and line graphs were clearly categorized into clusters whereas there were conflicts in categorizing illustrations and diagrams.

**Keywords:** categorization of graphics, card sorting

The present study examines how master students cluster different graphics into diverse categories. Graphics is described as the basic sign system to sort, comprehend and communicate information (Bertin, 2011). Graphics have gained more and more importance in pedagogical material and in text books. Irrespective of domains, use of graphical representations for teaching and learning is very common. Educational graphics are developed keeping in mind the level of the students and training of the teachers. To great extent, information is presented through diagrams, pictograms, maps and charts on paper and digitally.

## Different categories of graphical representations used in the study

Categorizing graphical representations is not a straightforward task. There are no predefined rules to distinguish graphical representations. We came across numerous definitions of each graphical representation but cartographers do not agree on one specific definition solely sufficient to define a particular graphical representation. Wilkinson and Wills (2005) mentioned that the notion of graphical representations is relatively undefined. We encounter many graphical representations while learning in schools and in everyday life but it is not obvious to distinguish them and even more difficult to interpret.

The cartographers spend a good deal of time in developing attractive, informative and concise graphical representations and wonder why people don't understand their creations. It's the fault of neither the cartographers nor the readers but it's the education and capability of understanding graphical representations. Gallimore (1990) more deliberately emphasizes the aspect that teaching of graphical representations should be focused and primary schools should teach graphicacy just like reading, writing and mathematics skills. In a study, Aldrich and Sheppard (2000) draw attention to the fact that graphicacy is hardly taught explicitly and it's frequently expected that students will comprehend the graphical representations and emphasize teaching graphic skills just like literacy and numeracy. Our main **questions** are: How do students categorize graphics? Do they follow a priori "expert" categorization? And finally, what are characteristics of difficult cases in categorizing?

## Method

*A priori categorization:* Bertin (2011) divided graphics into three groups. The first group, diagrams, includes bar graphs, pie charts, line graphs, diagrams etc. The second group, networks, comprises of flow charts, hierarchy charts, trees and charts of inclusive relationships. Maps constitute the third group. We developed definitions based on dictionaries and followed them in categorizing graphics. The categories were Map, Diagram, Schema, Line graph, Illustration, Pie chart, Flow Chart,

Pictogram, Bar graph, Table and Hierarchy graph. We selected 50 graphical representations from different newspapers, textbooks, with information about geography, physics, chemistry and mathematics etc. Graphics were selected in order to have a large variety of graphical categories described above. Headings and legends in all graphical representations were defined clearly for better understanding. Some of graphics did not have titles, so they were given a title to keep all the graphics likewise. The set of 50 graphics were categorized by two raters and Cohen's kappa test was computed to check the reliability. The resulting kappa of .67 indicated that there were certain differences in categorizing diagrams, schemas, illustrations and pictograms. We merged diagrams and schemas. Cohen's kappa test was recomputed which resulted in a value of .75. The remaining differences were discussed to inter-subjectively decide on final categorizations.

*Card sorting technique:* Twenty Master students in Education of the University of Grenoble II were divided into five groups. We used open card sorting method and asked the participants to categorize 50 graphics into groups based on similarities. The participants were asked to give a name to each of the groups. It took almost 20 minutes for the card sorting activity.

## **Results**

A hierarchical cluster analysis with SPSS using Ward's method produced seven clusters. We compared these seven clusters with the priori categorization.

Cluster 1 contained 13 diagrams/schemas, one illustration and a table.

Cluster 2 contained eight typical maps.

Cluster 3 contained five more atypical maps.

Cluster 4 contained six line graphs and one illustration (see right-hand side of Figure 1).

Cluster 5 contained three pie charts, two pictograms, a table and a bar graph.

Cluster 6 contained two diagrams, two flow charts and a hierarchy chart.

Cluster 7 contained two illustrations and a diagram.

Maps and line graphs were two categories where students were agree with the priori categorizations. Even though, the right-hand side of figure one shown below is an example of how students categorized the illustration as a line graph. Whereas maps was the one category closely followed by the students. On the other hand, students categorized most of the illustrations and pictograms as diagrams whereas they were categorized as illustrations or pictograms in priori categorization. Although, maps were categorized in multiple groups despite the fact that they were labelled as maps but the participants put them in different groups. The maps which presented some unusual information were not categorized with usual ones. The students were ambiguous while interpreting multivariate graphics and categorized some of them as illustration, diagrams and maps etc. Familiar graphics as Roth, Pozzer-Ardenghi, & Young (2005) mentioned in their study were easily identified as compare to unfamiliar ones which categorized in multiple groups.



Figure 1. Two atypical graphics

Figure 1 shows two atypical graphics. The one on the left was categorized as diagram/Schema in priori categorization but students categorized it as image, some as diagram and some as illustration. It was one of atypical types of graphic which created a lot of ambiguity and students discussed it a lot before categorizing in a particular group. Use of colours for different parts of the diagrams created a lot of ambiguity and it was considered difficult to compare different aspects as well. We categorized the one on the right as an illustration, but students grouped in the cluster with line graphs. To us, both the absence of axes, legends, etc. and the fact that the line probably not reflects an actual data set are arguments for categorizing it as an illustration. Students seem to cluster on superficial grounds.

## Conclusion

It was noticed in both priori and student categorization that the multivariate graphics were considered difficult to interpret. Teaching of graphics as an individual aspect is important so that student can identify, interpret and comprehend the graphics.

## References

- Bertin, J. (2011). *Semiology of Graphics: Diagrams, networks, maps*. Redlands: Esri Press.
- Gallimore, M. (1990). *Graphicacy in the Primary Curriculum*. Paper presented at ICOTS. (pp.140-143).
- Parmar, R. S., & Signer, B. R. (2005). Sources of error in constructing and interpreting graphs: a study of fourth- and fifth-grade students with LD. *Journal of Learning Disabilities*, 38, 3.
- Roth, W.-M., Pozzer-Ardenghi, L., & Young, H. J. (2005). *Critical graphicacy: Understanding visual representation practices in school science*. Dordrecht: Springer
- Wilkinson, L., & Wills, G. (2005). *The grammar of graphics*. New York: Springer.

# Mere Presence, Object Orientation and Perspective Taking

*Adriana Baltaretu*

*Alfons Maes*

*Carel van Wijk*

Tilburg Center for Cognition and Communication (TiCC)

P.O. Box 90153 NL-5000 LE Tilburg, The Netherlands

A.A.Baltaretu@uvt.nl, Maes@uvt.nl, chvwijk@uvt.nl

**Abstract:** In a production experiment, we investigated whether perspective taking behaviour is indeed influenced by the mere presence of a person, as claimed by Tversky and Hard (2009). In addition, we tested the effect of object orientation on spatial perspective taking. Results show that mere presence of a person does indeed affect perspective taking. Object orientation does have an effect: not on the linguistic form of the description but on the time to start a description. For objects pointing toward the speaker latencies are shorter than for objects pointing toward the addressee. None of the experimental factors affected the linguistic form of the description. Only one of them, object orientation, affected a temporal aspect of the utterance.

**Keywords:** spatial perspective taking, object orientation, location descriptions, addressee awareness

For any teacher giving instructions to a learner - either on paper, on the screen, or in face to face interaction - there is always the question which perspective to take: that of the learner, or that of the teacher. It might be confusing when teachers speak from their viewpoint, and easier if they take the learners' frame of reference. In the latter case, the teacher faces the task of Spatial Perspective Taking.

Spatial Perspective Taking is considered to be an alignment process, in which a speaker takes the perspective of somebody else (Kessler & Thomson, 2010). Spatial perspective taking is a basic form of reasoning about other people's representations of the world, and thus crucial for successful communication. In a recent study, Tversky and Hard (2009) had respondents describe spatial relations between two objects on a table in photographed scenes (see Figure 1). When the scene included a person, about a quarter of participants adopted the viewpoint of the person in the scene. In a second study, when the visual presence (a reaching person) was also combined with a linguistic cue (action verb questions) about half of the participants took the other's perspective. On the basis of these findings, Tversky and Hard (2009: 125) concluded that "the mere presence of another elicits spatial perspective taking." This conclusion, however, is not sound yet, as their stimuli contain other visual clues that might have promoted the other's perspective. First, the pictures show more than a 'mere' person, they show an addressee strongly involved in the situation (the addressee looks or looks and reaches at the book); second, the object to which the person is attending (the book in figure 1) is 'functionally' oriented towards the addressee. According to Gibson (1979) objects elicit affordances, possibilities for actions, which in this case apply in the first place to the addressee in the scene, which in turn may promote to take the addressee's perspective.



(a) Reaching

(b) Looking

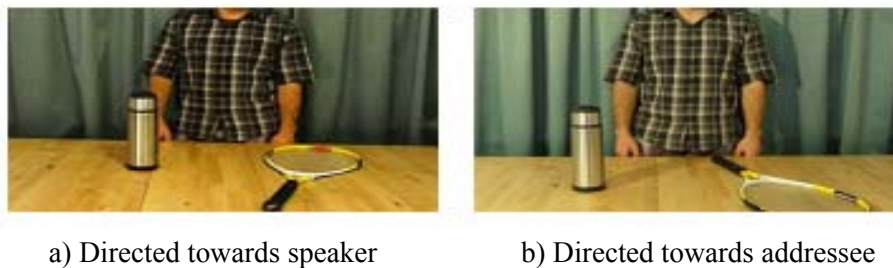
(c) No person

Figure 1 - Stimuli used by Tversky and Hard (2009)

In our study, we replicated Tversky and Hard's (2009) experiment taking into account these considerations. We asked respondents to describe the position of two objects with respect each other, in a scene with a person 'merely' present (not looking or reaching). We did not use any linguistic addressee cue. We varied systematically the functional orientation of objects (directed towards the addressee or to the speaker). Finally, instead of using one pictorial scene only, we used a large number of items.

### Materials

We constructed pictures depicting a table with two objects, and with a person standing behind the table whose head is not shown, and who is holding his arms next to his body (figure 2). Each picture consisted of a symmetrical object (e.g., a thermos bottle) and a directed object (an object with a functional orientation, e.g., tennis racket). The directed object was varied in two ways: position (to the left, right of the symmetrical object) and orientation (toward the speaker, towards the addressee). This resulted in 24 pictures, 12 for each orientation, 6 with the directed object to the left, 6 to the right. In the verbal instruction, we omitted the question format and simply asked participants to describe the position of the objects relative to each other. For each response two scores were derived: the linguistic form of the response (from speaker perspective or from addressee perspective), and the response latency (the time needed to start with the description).



*Figure 2* – Example of experimental stimuli: object with speaker (a) and addressee (b) orientation respectively

### Participants and procedure

Seventy-six Tilburg University undergraduate students (15 male, 61 female) participated in return for course credits. All participants were unaware of the experiment's goal.

Participants were placed in a sound proof cabin in front of a desktop computer with a headset microphone. They read the instructions from screen. Next, they carried out a practice trial (two pictures with 2 symmetrical objects). Then the twenty-four experimental trials were presented in one of two presentation orders: first 12 pictures with one object orientated towards the addressee (=A), followed by 12 pictures with one object directed towards the speaker (=S) or vice-versa (A/S versus S/A).

After producing a description, participants pressed the space bar to start the next trial, until they had completed all trials. Stimulus presentation and recording of reaction times were done using E-prime 2.0 software (Psychology software tools, Inc., 2007); the recording of the description was done using open source software (Audacity 1.3.14).

### Results

Due to technical problems, speech data of six participants could not be analyzed; they were excluded from all further analyses. We analyzed 70 participants' data (35 subjects A/S presentation order and 35 subjects in the S/A presentation order).

Each description has been scored for perspective taking in the following way. Situation (a) presented in Figure 2 can be described in four ways, see (1). Situation (b) in Figure 2 offers more descriptions, see (2). Responses e. and f. show that the speaker was aware of the presence of a potential addressee and has incorporated that person's perspective, that is, the speaker applied an addressee perspective. Out of 70 participants, the large majority (N=57) used their own (speaker) perspective. Thirteen participants consistently used an addressee perspective (7 in A/S; 6 in S/A). The repeated presentation of pictures with an object directed towards the addressee never made a participant change perspective. On the other hand, once they started with an addressee perspective, they used it consistently. On the basis of their responses the participants have been split into two perspective groups: speaker vs. addressee perspective.

- |  |  |
|--|--|
| <p>(1) The tennis racket is</p> <ul style="list-style-type: none"> <li>a. on the right</li> <li>b. on my right</li> <li>c. on the right side of the table</li> <li>d. to the right of the terms</li> </ul> | <p>(2) The tennis racket is any answer from a to d</p> <ul style="list-style-type: none"> <li>e. on the left side</li> <li>f. on my right, but to the left of the man</li> </ul> |
|--|--|

For each Orientation, the scores of the twelve pictures were averaged. These mean response latencies have been evaluated with a repeated measurement ANOVA with within factors, Orientation (toward speaker, toward addressee) and two between factors Presentation Order (A/S, S/A) and Perspective Taken (speaker, addressee). An interaction was emerging between Orientation and Perspective Taken ( $F(1,63)=3.19$ ,  $p=.08$ ,  $\eta_p^2=.048$ ). Split analyses were run for the two participant groups. The group that took the speaker perspective showed an effect of Orientation ( $F(1,53)=9.79$ ,  $p<.005$ ,  $\eta_p^2=.156$ ): they responded faster (137ms) when the object was directed toward them (1481 versus 1618 ms). The group that took an addressee perspective, showed the reversed pattern, faster when the object was directed toward the addressee (1729 versus 1764 ms) but this difference did not reach significance ( $F<1$ ).

## Conclusions

The results confirm that when people describe the location of objects in the nearby space, the mere presence of another person is strong enough to elicit perspective taking behaviour: about 80 percent of the speakers took their own position as point of departure, 20 percent took the other person into account as well, comparable to the findings in Tversky and Hard (2009). Once a perspective had been chosen, speakers used it consistently; no one changed perspective during the course of describing 12 subsequent pictures. The orientation of an object was not strong enough to change the addressee perspective. Yet, it influenced the speaker's performance: response latencies were shorter when the picture was in line with the perspective taken.

## References

- Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Kessler, K., & Thomson, L., A. (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. *Cognition*, 114 (1), pp.72-88.
- Tversky, B., & Hard, M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, 110(1), pp.124-129.

# A Comparison of Different Levels of Interactions when Using the Isolated-Elements Strategy

Dominique Bellec<sup>1</sup>, André Tricot<sup>1</sup>, Paul Ayres<sup>2</sup>

<sup>1</sup>University of Toulouse, France and <sup>2</sup>University of New South Wales, Sydney, Australia  
Dominique.Bellec@ac-poitiers.fr, Andre.Tricot@univ-tlse2.fr, p.ayres@unsw.edu.au

**Abstract.** Complex materials are difficult to learn. One solution is to reduce the complexity in an initial phase of learning. This effect is obtained when components are presented with no interactions between them, but also when presented with some interactions still present. In this experiment, forty-three participants had to learn about the effectiveness of hybrid engines as a solution to the reduction of greenhouse gas emissions. This material involved nine components or variables and a complex set of interactions. Participants were randomly assigned to three different versions of the material: isolated elements with no interactions were presented before the whole interacting system; partially interacting elements were presented before the whole interacting system; direct presentation of the whole system (twice). For recall questions all three groups performed the same, but on transfer questions the non-interacting group performed worse and expended greater mental effort, suggesting that greater exposure to interactions are required in complex domains.

**Keywords:** Complex learning, Isolating elements, Cognitive load

## Introduction

Many learners are required to learn about complex systems, phenomena or procedures. Such materials can be very difficult to learn. According to cognitive load theory element interactivity is a major construct underpinning the learning of complex materials (see Sweller, Ayres & Kalyuga, 2011; Sweller & Chandler, 1994). Materials high in element interactivity are difficult to learn because they require more working memory resources to deal with the multiple interactions between elements. A promising solution to this problem has been to break down the complexity of problems by isolating the interacting elements in an initial phase of learning. The *isolated interacting elements effect* is obtained when interacting elements are presented in isolation before presenting the full interacting material (e.g. Sweller, 2010). This procedure is superior to presenting the full interacting material twice (Pollock, Chandler & Sweller, 2002; Ayres, 2006). This effect has been compared or interpreted as a pre-training effect (Mayer, Mathias & Wetzell, 2002; Clarke, Ayres & Sweller, 2005), simple-to-complex sequencing (van Merriënboer, Kester & Paas, 2006), molar-modular effects (Gerjets, Scheiter & Catrambone, 2006), and the use of subgoals (Catrambone, 1998).

These different studies, referenced above, share one important feature: when learning a complex system, phenomena or procedure, it is more efficient to acquire knowledge about components before learning the whole, than trying to learn the whole directly initially. But our analyses of these studies suggest there is also a significant difference between them: some are obtained when smaller components are presented in isolated form with no interactions between them, whereas others are presented as components but with some interactions still present. Hence this difference is worthy of further investigation.

The aim of this study was to compare these two alternatives to a direct presentation of the whole system. In a two-stage design the following three strategies were compared: a) isolated elements with no interactions were presented (stage 1) before the whole interacting system (stage 2); b) partially interacting elements were presented before the whole interacting system; c) the whole interacting system was presented twice. This design would test if the isolated interacting elements effect can be considered as a general effect of simplification of the material per se, or if it is more closely linked to the isolation of interacting elements rather than simplification. The study was conducted with instructional materials that feature both texts and graphics.

## Method

Forty-three students of a technical college from Poitiers, France, participated in this experiment. They had to learn about the effectiveness of hybrid engines as a solution to the reduction of greenhouse gas emissions. This computer-based material involved nine components or variables: strength, power, torque, energy, consumption, transformation of energy (mechanical - electrical), kinetic energy, electric energy storage (battery), production of carbon by gas. The goal was to learn the complex set of interactions between these components and variables in four different driving situations (start, average speed, hard acceleration and braking). Three versions of the material were designed to represent three strategies:

a) The Non Interacting Strategy: Each component or variable with no interaction with other components or variables (9 pages) was presented initially (Stage 1). Each page contained one text, one picture or graphic and the measuring unit (when the element was a variable). This was followed by a presentation (Stage 2) of the whole interacting system (11 pages). The two stages are never available at the same time: the participants can go forward and backward within the stage 1, when they decide that they go to stage 2, then it is not possible to go back to stage 1.

b) The Partially Interacting Strategy: Each component or variable with interactions with one other variable or component was presented initially in Stage 1 (9 pages).. For example, the relation between strength and power (and speed) was presented in the second page, describing power. Each page contained one text, one picture or graphic and the measuring unit (when the element was a variable). This was followed by a presentation (Stage 2) of the whole interacting system (11 pages).

c) The Total Interacting Strategy: The whole interacting system was presented twice (11 pages then the same 11 pages).

Each group was randomly assigned to one version of the materials. A pre-test (29 questions) was used, one week before the experiment, to evaluate the students' prior knowledge. To investigate if different types of questions influenced the effectiveness of the strategies 20 questions were designed as recall questions, as the answers were explicitly contained within the materials; and 9 questions were designed as transfer questions, where the answers were not explicitly but stated but had to be inferred. During the instructional phase the participants were free to spend as much time as they wanted on each learning phase. They were also free to go forward and backward within the two set of materials corresponding to the first and the second learning phase. The number of pages studied and the time spent on each page was recorded.

A post-test (the same 29 questions as the pre-test) was assigned at the end of the second phase. A subjective measure of mental effort (see Paas & van Merriënboer, 1994) using a 9-point Likert scale was completed after both learning phases.

## Results

Performance scores were analysed by ANOVA using improvement scores for each participant (Post-test – Pre-test). There was no significant difference between the three conditions for recall measures:  $F(2,40) = 1.79$ , ns. For the transfer questions, there was a significant difference:  $F(2, 40) = 3.49$ ,  $p < .05$ , partial  $\eta^2 = 0.15$  (large effect). Tukey B post hoc tests revealed one significant comparison in that the Total-interacting system was superior to the Non-Interacting system. For mental effort during phase 1, there was no significant difference  $F(2,40) = 1.47$ , ns. However there was a significant difference for Phase 2:  $F(2, 40) = 5.25$ ,  $p < .01$ , partial  $\eta^2 = 0.21$  (large effect). Tukey B post hoc tests revealed the Non-Interacting group invested more mental effort than the other two interacting groups.



Table 1: Performances, mental effort and time spent on three different conditions.

	Recall questions		Transfer questions		Mental Effort	
	PreTest	PostTest	PreTest	PostTest	Phase1	Phase2
Non Interacting	11.6	13.4	3.1	3.3	4.00	6.44
Partially Interacting	12.0	13.4	2.0	3.7	4.50	4.44
Total Interacting	10.4	13.6	1.9	4.1	4.50	4.57

## Discussion

Unexpectedly, compared with previous research (see Pollock et al., 2002), isolating elements in the initial stage of learning about a complex system did not lead to better learning outcomes than initially showing interacting elements. Although there was no difference on recall problems, the non-interacting strategy was inferior on transfer problems, even though more mental effort was spent. These results suggest that more exposure of the various interactions may be needed to learn these materials successfully, particularly when learning very complex tasks. It is notable that previous research into isolated elements has featured less complex domains, and arguably this may have made the difference.

## References

- Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology*, 20, 287-298.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, 127, 355-376.
- Clarke, T., Ayres, P., & Sweller, J. (2005). The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. *Educational Technology, Research and Development*, 53, 15-24.
- Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, 16, 104-121.
- Mayer, R. E., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied*, 8, 147-154.
- Paas, F. G. W. C., & van Merriënboer, J. J. G., (1994) Measurement of cognitive load in instructional research, *Perceptual & Motor Skills*, 79, 419-430.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12, 61-86.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22, 123-138.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., & Chandler, P. (1994) Why some material is difficult to learn. *Cognition & Instruction*, 12, 185-233.
- van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology*, 20, 343-352.

# Using 3D Animation for Learning Functional Anatomy

*Sandra Berney<sup>1</sup>, Gaëlle Molinari<sup>2</sup>, Nady Hoyek<sup>3</sup>, Mireille Bétrancourt<sup>1</sup>*

<sup>1</sup>University of Geneva – TECFA, 40, Bld du Pont d'Arve, 1205 Geneva, Switzerland

<sup>2</sup>Distance Learning University Switzerland, Techno-Pôle 5, 3960 Sierre, Switzerland

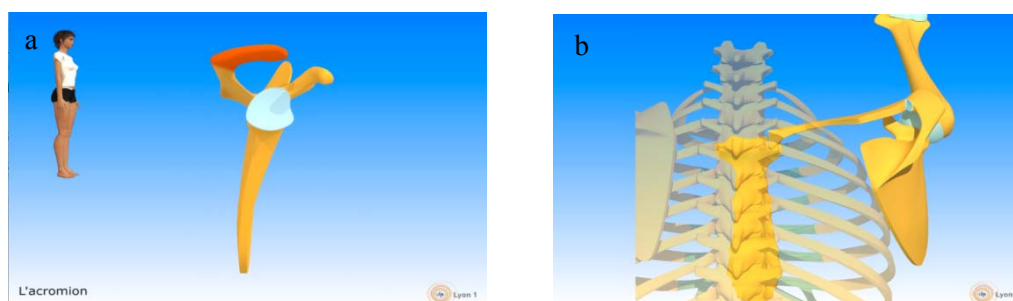
<sup>3</sup>CRIS, Université Lyon 1, 43 bd du 11 novembre 1918, 69622 Villeurbanne Cedex, France  
sandra.berney@unige.ch

**Abstract.** Learning functional anatomy requires the building of a dynamic mental representation to understand the structure and its behaviour. Animation, often used to represent dynamic systems, can also be used to depict the configuration of a 3D structure, as it provides direct visualization of change across the viewpoints. This paper reports a study comparing two versions of a visual instructional material (animated or static) in learning the structure and behaviour of the scapula. Results showed no effect of the conditions on performances, though locally the animation group was more accurate in performing some configuration tasks. Moreover, visuo-spatial abilities affected the performance but the interaction with the instructional version depended on the task.

**Keywords:** dynamic visualization, spatial abilities, anatomy learning, 3D representation

## Introduction

Learning functional anatomy, from textbooks and anatomical charts, requires the learner to mentally manipulate the anatomical structure to imagine its spatial orientation to gain further understanding of its dynamic behaviour. The building of such an accurate mental representation depends heavily on learner's spatial abilities, as they have to (re)create a dynamic mental model (Stull, Hegarty & Mayer, 2009). External visualizations, such as 3D animations, may bring an adequate solution to fill in the spatial difficulties encountered with static learning (Guillot, Champely, Batier, Thiriet & Collet, 2007; Hoyek et al., 2009), though they may be hard for novices to process. However, learning with external visualizations does not necessarily lead to better understanding, partially because of the intricate interplay of visuo-spatial abilities when learning (Hegarty, Kriz & Cate, 2003).



*Figure 1.* a) structure configuration: the scapula and its acromion process (in orange) in a 60° Y view and b) structure behaviour/movement: superior view of the scapula movement during shoulder flexion

Animated visualizations can serve various instructional functions, such as a) conveying the configuration of a system or structure (Figure 1a) and b) the structure behaviour or its movement (Figure 1b). The present research investigates the effect of two learning conditions (animated versus static visualizations) on the building of mental representations of structure and movement when learning functional anatomy of the scapula. We assumed that animation could help understanding the scapula behaviour by directly showing the movement. Animation could also support the construction of a mental representation of the scapula structure within the anatomical 3D space of the body by providing transitions between the viewing perspectives.

## Method

### *Participants and Design*

Forty-nine 1<sup>st</sup> year students enrolled in the physical education degree at the University of Lyon 1, France, voluntarily participated in the study. They were randomly assigned to one of two learning conditions: a static visualization ( $n=27$ ) and an animated visualization conditions ( $n=22$ ).

### *Instruction material*

The learning material developed by Icap Université Lyon 1 consisted of two 3D animations: a) the structure of the scapula and 6 of its features (acromion process, inferior angle, coracoid process, lateral border, spine, neck), and b) the scapula shoulder flexion movement, that is when the arm is moving from the standard position (arm along the body) upward to the front. In both animations, four anatomical orientation views were sequentially presented:  $0^\circ$  view (scapula posterior surface),  $60^\circ$  view (Y view of lateral scapula),  $180^\circ$  view (anterior surface), and *superior view*. In addition, in the structure material, a small character acted as a permanent spatial anatomical reference, whose orientation changed according to the viewpoints. Two versions of these materials were designed; a) an animated visualization version and b) a static visualization version, which presented simultaneously the main small-scale images of each views. Both learning conditions materials had the same duration.

### *Knowledge tests*

Students' learning was assessed with 3 *structure* and 2 *movement* tasks. The *structure* tasks consisted of 1) a **feature identification** task assessing the recall of features' name and anatomical relative location; 2) a **rotation of the scapula** task assessing the understanding or recognition of various scapula rotations within the anatomical space; and 3) an **orientation reference** task for the recognition of the scapula position with regard to the orientation reference character. *Movement* tasks involved 4) a **movement identification** task for the recognition of dynamic excerpts of dis/similar phases of the shoulder flexion movement; and 5) a **movement order** task assessing the understanding of the scapula movement during shoulder flexion by ordering five static images of different motion states.

### *Procedure*

Participants had an initial 2-minute study phase of scapula general information (a 86-word text coupled with a scapula labelled picture), followed by the instructional *structure* material and its 3 tasks. Then, the *movement* material was presented followed by its 2 tasks. All learning material was presented twice. Additional cognitive measures assessed mental rotation abilities (MRT, Vandenberg & Kuse, 1978) and dependence towards the field (GEFT, Oltman, Raskin, & Witkin, 1971), from which high and low groups were defined by a quick cluster analysis.

### *Results*

Findings revealed no overall advantage when learning with animated visualizations. Performance of the 2 groups did not differ on *structure* (Wilks'  $\lambda$  ( $\lambda$ ) = .91,  $F(3, 45) = 1.34$ , *n.s*) nor on *movement* tasks ( $\lambda = .97$ ,  $F < 1$ , *n.s*). However, significant global effects of mental rotation abilities (MRT) were found on the overall *structure* ( $\lambda = .69$ ,  $F(3, 43) = 6.23$ ,  $p = .001$ ), and *movement* performances ( $\lambda = .82$ ,  $F(2, 44) = 4.61$ ,  $p = .015$ ). Moreover, low MRT students performed better on the *feature identification* task when learning with animation ( $M = 24.07$ ) compared to static learning ( $M = 20.11$ ), whereas the reverse pattern was found for high MRT students ( $M_{static} = 28.10$ ;  $M_{animation} = 25.22$ ;  $F(1, 45) = 5.46$ ,  $p = .024$ , partial  $\eta^2 = .10$ ). No effect of field dependency (GEFT) and interaction effect were found.

*Accurate items.* A significant interaction between *structure* tasks and MRT clustering showed a quadratic trend ( $F(1, 45) = 5.71, p = .021, \eta^2 = .11$ ). *Feature identification* and *orientation reference* tasks had an unusual but specific pattern (low MRT scores > high MRT scores), whereas the *scapula rotation* task presented an opposite trend (low MRT scores < high MRT scores).

*Error items:* Compared to the static learning group, students in the animated learning condition made fewer errors, when a) recalling the neck characteristic in the *feature identification* task ( $M_{\text{animation}} = 1.31; M_{\text{static}} = 2.33; F(1, 47) = 4.35, p = .042, \eta^2 = .53$ ), b) comparing previewed canonical images in the *scapula rotation* task ( $M_{\text{animation}} = -.37; M_{\text{static}} = .30; F(1, 47) = 6.23, p = .016$ , partial  $\eta^2 = .11$ ), and c) comparing items differentiated by a 90° vertical rotation in the *scapula rotation* task ( $M_{\text{animation}} = -.33; M_{\text{static}} = .27; F(1, 47) = 4.84, p = .033, \eta^2 = .09$ ).

## Discussion and future research

Overall, findings suggest no differences in performances across the learning conditions concerning the building of a mental representation of a 3D structure, as learners from both conditions performed equally on the *structure* and *movement* tasks. Explanations might be that a) the mandatory 2-minute learning phase upset the follow-up learning, b) information from both instructional materials was equivalent and/or c) tasks were not competitive enough to reveal differentiated performances. In line with the literature, we found that the building of mental representation of a 3D structure was largely influenced by learners' spatial abilities, particularly mental rotation abilities. The influence of the mental rotation levels on the *structure* performances differed across the tasks, suggesting that the tasks might not all rely on the same spatial abilities. The follow-up analyses revealed an interesting result regarding rotations. The animation group made fewer errors of rotation judgment with items where the scapula's view was 90° vertically rotated. Whereas both learning groups watched the same anatomical views, the difference lies in the motion. The animated learning group watched the scapula horizontally turn from 0° to 180° and then vertically to the superior view, while the static group saw sequentially the 4 frames. This result may suggest that the *structure* animation could help learners build the mental representation of the 3D scapula. Complementary research is needed to assess this hypothesis.

## References

- Guillot, A., Champely, S., Batier, C., Thiriet, P., & Collet, C. (2007). Relationship between spatial abilities, mental rotation and functional anatomy learning. *Advances in health sciences education: theory and practice*, 12(4), 491-507. doi:10.1007/s10459-006-9021-7
- Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, 21(4), 325-360.
- Hoyek, N., Collet, C., Rastello, O., Fargier, P., Thiriet, P., & Guillot, A. (2009). Enhancement of mental rotation abilities and its effect on anatomy learning. *Teaching and learning in medicine*, 21(3), 201-6. doi:10.1080/10401330903014178.
- Stull, A. T., Hegarty, M., & Mayer, R. E. (2009). Anatomy Learning with Virtual Objects. *Intelligence*, 39-44.

# Graphicacy: Do Readers of Science Textbooks Need It?

Mireille Bétrancourt, Shaaron Ainsworth, Erica de Vries, Jean-Michel Boucheix & Richard K. Lowe

TECFA, University of Geneva, Switzerland, Mireille.Betrancourt@unige.ch

University of Nottingham, United Kingdom, Shaaron.Ainsworth@nottingham.ac.uk

LSE, University of Grenoble II, France, Erica.deVries@upmf-grenoble.fr

LEAD-CNRS, University of Burgundy, France, Jean-Michel.Boucheix@u-bourgogne.fr

Curtin University, Perth, WA, Australia, R.K.Lowe@curtin.edu.au

**Abstract.** In contrast to literacy and numeracy, graphicacy is under-explored. Yet it is not only linked to fundamental cognitive abilities, but is also important in all areas of instruction. This paper reports an analysis of the graphics that appear in K-12 science school books. The questions were the following: what skills might school students need to understand these graphics? Do books provide guidance or instructions to help students process these graphics? We found that many graphics required a high level of graphicacy, in terms of strategies and skills required to understand them. Moreover, strategic guidance or instructional help were rarely provided.

**Keywords:** graphic representations, graphicacy, textbooks, interpretative support

## Introduction

Several kinds of communicative competencies have been identified that refer to abilities, beyond domain-specific knowledge, to use knowledge and skills “to achieve one’s goal in various situations (personal, professional, civic) in order to participate effectively in society” (OECD). *Graphicacy* is a graphic-specific communicative competence whose importance has hitherto been underestimated. It can be generally defined as the abilities to *interpret* and *generate* graphical representations, such as charts, diagrams, maps, and graphs (de Vries & Lowe, 2010). These skills and knowledge are used in learning or performance in contextualised tasks such as using a map to navigate an unfamiliar town, interpreting a diagram to solve a paper jam in a printer, or to construct a graph that presents data for an academic assignment. In this paper, we focus only on the interpretative aspects of graphicacy.

Understanding a graphical representation requires grasping its component elements as well as their global organization and relating the graphic to domain knowledge. Whereas text conventions function at a general level (except for technical vocabulary), conventions in graphics are far more diverse and domain-dependent. Moreover, different types of graphics have their own distinctive and very specific demands. With text, various agreed classifications are used that define functionally different types of representation and are identified as distinct genres (see de Vries and Lowe, 2010). In contrast, there is no such agreement on the number and types of graphics genres, and no clear-cut universal graphic conventions. Furthermore, although children do seem to develop incidentally some capacity to select, understand and construct graphics, what diSessa (2004) called meta-representational competence, this is not sufficient given society’s increasing reliance on graphics. It must be supplemented by the use of appropriate strategies, which also depends on the domain and graphic specific knowledge.

Our work on graphicacy began with a concern about the graphicacy demands typically faced in everyday school situations. Consequently, we conducted a preliminary examination of graphicacy requirements for interpreting K-12 science school books from five countries. Our initial exploratory questions were as follows: what skills and knowledge are required from learners to understand the graphic representations provided? Do books provide guidance on how to process these graphics?

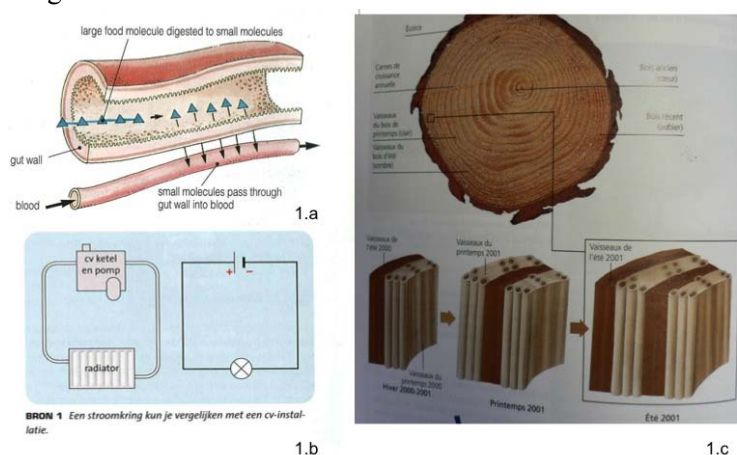
## Method

Graphic representations were extracted from five textbooks (see references) from different countries (Australia, UK (in English), France, Switzerland (in French) and Netherlands (in Dutch)), all used in upper primary school by children from 10 to 12 years old. These materials were selected because the target audience is relative novices in the domain, and because the books abound with graphical representations. In order to obtain a broad sample, we selected both typical and atypical instances of

visual representations with regards to existing features extracted from the semiotic and cognitive literature. Alternative approaches such as selecting a single topic and identifying all representations associated with it or identifying representations by existing classification schemes (e.g. Levin, Anglin & Carney, 2007; Ainsworth, 2006) were rejected as not appropriate for identifying generic aspects of graphicacy. The graphics from all five books were successively analysed in terms of processes and knowledge considered necessary to understand them, and whether interpretative guidance was provided by the textbook. Their relation to the accompanying text, although an important question, was not considered here, except for the captions and instructions (mostly questions to answer from the graphics). Figure 1 presents some of the graphics.

## Analysis

The first observation is that there is a large variety of graphic types from simple line drawings to complicated perspective portrayals, using various styles and codes (e.g. line graphs, histograms, pie charts, time-lapse photographs, cross sectional diagrams, photographs, cartoons, networks, drawings). This diversity might be problematic since learners have to learn to interpret graphics whilst simultaneously learning content.



Such combinations were used *inter alia* to show different periods of time (e.g. steps of a process, or evolution over time, Figure 1c), or to represent the same content but in different ways (e.g. realistic vs. abstract depiction of the same object Figure 1b), or to show a detailed aspect as well as global overview (Figure 1c). Consequently, in order to grasp the intended meaning, the graphic learner must have the knowledge and skills described above and must identify the similarity and the differences between the representations in order to deduce the semantic relations involved relative to the topic at hand. One of the most common types of combinations found in our analysis was ‘paired graphics’ where two related but different depictions are presented together (e.g. Figure 1b). We therefore conducted a preliminary empirical study of how such combinations are processed (Boucheix et al, submitted).

## Conclusion

The textbooks we examined present many graphics that are complex for learners to interpret; yet instructions on how to read the graphics were rarely provided in the books and prior research suggests that teachers often neither recognize the need for this instruction nor know how to succeed in doing so (Eilam, 2012). We argue that graphicacy has a distinct constellation of capacities that future work should investigate using multiple methods and approaches. Furthermore, because interpretative guidance is so rarely given in textbooks, it needs specific attention in schools and so increasing attention must also be paid to how to support teachers to gain graphicacy knowledge themselves.

## References

- Ainsworth, S. E. (2006). Deft: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198.
- Aldrich, F. K., & Sheppard, L. (2000). Graphicacy '': the fourth ' R '? *Primary Science Review*, 8 - 11.
- Boucheix, J.M, Lowe, R. K., Ainsworth, S., Bétrancourt, M., & de Vries, E. (submitted). Paired graphics: An exploratory study of graphicacy. Paper submitted at the *EARLI SIG 2 Conference in Comprehension of Text and Graphics meeting*, Grenoble, France.
- De Vries, E., & Lowe, R. K. (2010). Graphicacy: What does the learner bring to a graphic? Paper presented at the *EARLI SIG 2 Conference in Comprehension of Text and Graphics meeting*, Tübingen, Germany.
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22(3), 293-331.
- Eilam, B. (2012). Teaching, learning, and visual literacy: The dual role of visual representation in the teaching profession. New York, NY: Cambridge University Press.
- Hegarty, M. (2010). Components of Spatial Intelligence. In B. H. Ross (Ed.) *The Psychology of Learning and Motivation* (pp. 265-297), Vol 52. San Diego: Academic Press.
- Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration: I. Basic research* (pp. 51-85). New York: Springer.

## Textbooks

- Spotlight Science 8 (2001) Nelson Thornes; Exploring Science 7 (2008) Pearson Schools.
- Science Aspects 1. (2009). G. Linstead, O. Gyder, G. Przywolnik, L. Salfinger, T. Herbert. Pearson Heinemann
- Sciences, Cycle 3. (2003). , J.M. Rolando, G. Simonin, P. Pommier, J. Nomblot, J.F. Laslaz, S. Combaluzier. Magnard
- Sciences et technologies. Cycle 3 (2002). Guichard, J., Hebrard, J. (dir.), Zana, B. (2002). Hachette Education, Paris.
- Pulsar, NASK, havo/vwo. (2006). Van den Berg, A., et al. Noordhoff Uitgevers, Groningen, The Netherlands.

# The Influence of Annotation in Graphical Organizers

*Eniko Bezdan, Liesbeth Kester, Paul Kirschner*

Open University of the Netherlands, Centre for Learning Sciences and Technologies

P.O. Box 2960, 6401 DL Heerlen, The Netherlands

eniko.bezdan@ou.nl, liesbeth.kester@ou.nl, paul.kirschner@ou.nl

**Abstract.** The effect of annotation in graphical organizers of hypertext learning environments was examined on learning outcomes of low and high prior-knowledge learners for knowledge acquisition and comprehension. Two hierarchical and two network organizers – with or without annotation - were compared. The annotation described the nature of the conceptual relationships depicted in the organizers. No differences were found for acquiring factual knowledge. For comprehension, the annotated hierarchical organizer negatively impacted low prior-knowledge learners as compared to a non-annotated hierarchical one. When the two groups were compared, high prior-knowledge learners scored significantly higher than low prior-knowledge learners with an annotated hierarchical organizer. In the non-annotated network condition, high prior-knowledge learners scored higher than low prior-knowledge learners. However, with an annotated network organizer, low and high prior-knowledge learners showed similar learning outcomes.

**Keywords:** hypertext, graphical organizer, annotation, prior-knowledge

## Introduction

Hypertext studies focusing on whether linear, hierarchical, or network/semantic graphical organizers can counteract disorientation and help or hinder learning have yielded mixed results. Hierarchical graphical organizers have repeatedly been shown to have beneficial effects on the learning outcomes of low prior-knowledge learners in comparison with linearly or semantically structured organizers. In contrast, the effect of graphical organizers on learning outcomes of high prior-knowledge learners has at best been inconclusive. Most studies report no differences in learning outcomes for advanced learners irrespective of the graphical organizer offered to them (DeStefano & LeFevre, 2007).

The defining property of any hierarchical structure is that its information is organized in a top-down manner, from general to specific, irrespective of content (Shin, Schallert, & Savenye, 1994). In comparison, a network graphical organizer provides less structural guidance, as a network represents a much less obvious structure (Amadiou, Van Gog, Paas, Tricot, & Mariné, 2009). In a network, any node can be connected to any other node, with the consequence being that the conceptual relationships have to be inferred by the learner. Learning with a network organizer, therefore, requires additional processing which can overburden the cognitive capacities of novices (DeStefano & LeFevre, 2007). For high prior-knowledge learners, this poses less of a problem, as they can consult existing cognitive schemata in their long-term memory to make inferences from ambiguous conceptual relationships. Also, it is often expected that network organizers will benefit high prior-knowledge learners, as a network represents a more complex representation of information than hierarchy. Potelle and Rouet (2003), for example, hypothesised that the categorical knowledge captured in hierarchical structures cannot provide new information to high prior-knowledge learners, whereas a network map displaying causal relationships does so by increasing connectivity within their existing mental models. Contrary to their expectation, Potelle and Rouet found no significant differences for the effect graphic organizers on learning outcomes of advanced learners.

The ambiguity associated with network organizers can be reduced by explaining the depicted conceptual relationships in them. The present study examined the extent to which explicitly defining the relationships between concepts - in the form of annotations in a hierarchical or network organizer - differentially affects low and high prior-knowledge learners. It was hypothesised that a annotations in a network organizer should alleviate its disadvantages for low prior-knowledge learners in comparison with high prior-knowledge learners. In contrast, adding annotations to a hierarchy was not expected to significantly influence learning outcomes, irrespective of prior knowledge. In line with earlier research



findings, high prior-knowledge learners were expected to show no significant differences across conditions.

## Method

The participants in this experiment were first year university students (mean age = 19.49,  $SD = 2.83$ ; 4 males and 153 females) studying the influence of Cognitivism on Learning and Instructional Science.

The four hypertext environments used in the four conditions of the experiment differed in the structure of the graphical organizer (i.e., hierarchical, network) and in the presence or absence of annotation (i.e., annotated, non-annotated). In one hierarchical and one network organizer, annotations were placed above the lines drawn between concepts, describing the conceptual relationships depicted. The network organizers retained all relationships of the hierarchical organizers, but a number of new, cross-referential relationships were added. Navigation in all conditions was only possible from the clickable concepts in the graphical organizer to corresponding text nodes and back.

Fourteen factual and fourteen comprehension multiple-choice questions were used to assess participants' knowledge before and after the learning phase. The multiple-choice questions had four answering possibilities, and were based on the content of the hypertext learning environment. Answering factual questions did not require knowledge of the content at related text nodes. Finding the right answer for a comprehension question required relating the information presented at the text node on which it was primarily based with the information presented at one or more related text nodes.

## Preliminary Results

Separate three-way ANOVAs were conducted, with the factors Structure (hierarchical, network), Annotation (no, yes) and Prior-Knowledge (low, high), for factual and comprehension questions respectively. Participants were categorized as belonging to the low and high prior-knowledge group by applying the median split of the pre-test scores.

No significant main or interaction effects were found for factual knowledge. There was a significant three-way interaction for comprehension between Structure, Annotation and Prior-Knowledge  $F(1, 149) = 5.56$ ,  $MSE = 5.17$ ,  $p = .020$ ,  $\eta^2 = .036$ . Simple effects analysis revealed that low prior-knowledge learners scored significantly higher with a non-annotated hierarchical organizer compared to the non-annotated network organizer  $F(1, 155) = 6.63$ ,  $MSE = 5.38$ ,  $p = .02$ ,  $\eta^2 = .041$  ( $M = 8.81$ ;  $SD = 2.46$  vs.  $M = 7.05$ ;  $SD = 1.85$ ). Furthermore, the low prior-knowledge group scored significantly higher with the non-annotated hierarchical than with the annotated hierarchical organizer  $F(1, 155) = 5.37$ ;  $MSE = 5.37$ ,  $p = .019$ ,  $\eta^2 = .033$  ( $M = 8.81$ ;  $SD = 2.46$  vs.  $M = 7.00$ ;  $SD = 2.38$ ). When comparing the two groups of learners, the high prior-knowledge group scored significantly higher with the annotated hierarchical organizer than low prior knowledge learners  $F(1, 155) = 6.22$ ,  $MSE = 5.35$ ;  $p = .014$ ,  $\eta^2 = .039$  ( $M = 8.84$ ;  $SD = 2.09$  vs.  $M = 7.00$ ;  $SD = 2.38$ ). High prior-knowledge learners also scored higher than low prior-knowledge learners with the non-annotated network organizer ( $M = 8.63$ ;  $SD = 2.56$  vs.  $M = 7.05$ ;  $SD = 1.85$ ), and this difference was significant  $F(1, 155) = 4.56$ ,  $MSE = 5.40$ ;  $p = .034$ ,  $\eta^2 = .029$ . The difference between the scores of the two groups of learners in the annotated network condition was not significant  $F(1, 55) = 0.04$ ,  $MSE = 5.56$ ;  $p = .840$  ( $M = 7.54$ ;  $SD = 2.28$  vs.  $M = 7.50$ ;  $SD = 2.50$ ).

## Discussion

As expected, low prior-knowledge learners reached better comprehension in the non-annotated hierarchical condition than in the non-annotated network condition. Similarly, the finding that high prior-knowledge learners did not show significant differences across conditions is in line with earlier findings in the literature. However, the annotated hierarchical organizer had a negative impact on the scores of low prior-knowledge learners, compared with the non-annotated hierarchical organizer.

When the two groups were compared, high prior-knowledge learners scored significantly higher than low prior-knowledge learners with the annotated hierarchical organizer. The mental integration of information communicated by annotations in a hierarchy on the one hand and the information at text nodes on the other may have required the allocation of added cognitive resources from low prior-knowledge learners. Descriptions of the relationships between text nodes (i.e., annotations) in the organizer were a repetition of descriptions of the same relationships at text nodes, and were, thus, possibly redundant. Dealing with redundancy is known to require the allocation of added cognitive resources.

In the annotated network condition, low and high prior-knowledge learners reached similar comprehension levels, whereas high prior-knowledge learners scored significantly higher for comprehension in the non-annotated network condition. Comprehension for the low prior knowledge group was higher in the annotated than in the non-annotated network condition, but this effect was not significant. It should be noted, however, that the annotated network organizer still scored lower for low prior-knowledge learners than the non-annotated hierarchical organizer. Although this difference was not significant, a word of caution is appropriate here; an annotated network organizer might be a good alternative for low prior knowledge learners, but only if the use of a hierarchical organizer is not appropriate or practically feasible.

### **Acknowledgement**

The authors would like to thank Professor Dr. Martin Valcke for his contribution to this study, and the first year students of Educational Sciences of Ghent University for their participation.

### **References**

- Amadiou, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction, 19*, 376-386.
- DeStefano, D., & LeFevre, J. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior, 23*, 1616-1641.
- Potelle, H., & Rouet, J. (2003). Effects of content representation and reader's prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies, 58*, 327-345.
- Shin, E., Schallert, D., & Savenye, C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in hypertext environments. *Educational Technology Research and Development, 42*, 33-46.

# Paired Graphics: An Exploratory Study of Graphicacy

*Jean-Michel Boucheix, Richard K. Lowe, Shaaron Ainsworth, Mireille Bétrancourt, Erica de Vries*

LEAD-CNRS, University of Burgundy, France, Jean-Michel.Boucheix@u-bourgogne.fr

Curtin University, Perth, WA, Australia, R.K.Lowe@curtin.edu.au

University of Nottingham, United Kingdom, Shaaron.Ainsworth@nottingham.ac.uk

TECFA, University of Geneva, Switzerland, Mireille.Betrancourt@unige.ch

LSE, University of Grenoble II, France, Erica.deVries@upmf-grenoble.fr

**Abstract.** Adults viewed 37 varied paired graphics derived from school science textbooks and explained meaningful aspects of each item. Verbal responses and eye tracking results indicated that participants understood most of the items and directed their attention to high relevance aspects of the graphics. Although some variables, such as the typicality of the graphic pairs appeared to influence comprehension, others such as content complexity seemed to have no effect. The findings will be used as baseline data for current studies of children's approaches to paired graphics and more formal investigations of the nature of graphicacy.

**Keywords:** graphicacy, science textbooks, paired graphics, comprehension, attention direction

## Introduction

Literacy and numeracy have long been regarded as fundamental capacities underpinning all aspects of education. However, with the rapidly increasing reliance on graphic forms of representation in educational materials, graphicacy has emerged as another fundamental capacity that learners need to develop (Balchin, 1985; Roth, Pozzer-Ardenghi, & Han, 2005). Graphicacy can be broadly defined as a set of capacities concerned with interpreting and generating information in graphic form (de Vries & Lowe, 2010). While there has been extensive empirical research into the nature of literacy and numeracy as well as their acquisition, this is not the case for graphicacy (Postigo & Pozo, 2004).

The study reported in this paper is an exploratory investigation of the paired graphics format used in many educational illustrations. Paired graphics were chosen for this initial investigation of interpretative graphicacy because they are both widely used and appear to require that learners engage in distinctive approaches to processing. A survey of mid-level school science textbooks showed that such graphics were common across both a range of countries and types of scientific content (see Bétrancourt, Ainsworth, de Vries, Boucheix & Lowe, 2012). Further, this paired format was used for a variety of instructional purposes, from showing the relationship between realistic and abstract depictions to portraying the 'before and after' states of target subject matter. Paired graphics typically use a standard layout formula of two pictures positioned horizontally adjacent on the display. This structural convention provides a signal that the component pictures are related and so should be processed together rather than independently. In order to deal effectively with paired graphics of even moderate complexity, it is likely that the learner would need to perform not only *within* picture processes but also *between* picture processes (cf. Winn, 1993). Further, this latter processing would probably require repeated inter-picture alternations of attention and search in order to extract the relational information present. For example, successive processing cycles could be used to establish the similarities and differences in the pair of pictures. In some graphic pairs, the left-to-right ordering is also an important interpretative cue (as with before-after depictions). Because interpretation is shaped by both the raw perceptual attributes of depictions and the learner's background knowledge, we conceive graphicacy as involving both bottom-up and top-down aspects.

To some extent, such processing approaches should be generic in that they could be applied across a broad range of different paired graphics examples. However, there would also be task-specific aspects that can be embedded within the generic approaches. We assume that these approaches would be acquired over time. However, in contrast to text, graphicacy is not currently the subject of explicit tuition and so is typically acquired incidentally. Adults with high levels of graphicacy should exhibit appropriate processing approaches when confronted with paired graphics, irrespective of the particular

content and specific purpose of the pairs presented. Their processing behaviour is likely to be very different from that of children whose graphicacy capacities are yet to develop. The present study used eye tracking to explore how adults processed a varied assortment of paired graphics. It was also designed to identify possible variables that could indicate the nature of graphicacy and how it might be assessed. We therefore used authentic examples from a variety of widely used science textbooks (rather than specially designed research materials). Data obtained will provide a baseline for current investigations of how young children deal with this type representation.

## Method

18 French university students (Psychology and Education, mean age = 37 years; SD = 11.24) participated for course credit. The 37 paired graphics items employed depicted a variety of topics and were based on actual examples from science school books for 10-12 year old children from five countries (Australia, England, France, Netherlands and Switzerland, Figure 1). Source materials for these items were chosen to provide a representative sampling from the different categories of paired graphics previously developed by the authors. The materials were re-designed for presentation without accompanying text (except for necessary labels). Participants were tested individually and instructed to imagine they were in a classroom situation using books with graphics. They were asked to study each of the items as they typically would in this situation and state what was most important to understanding the item's meaning. No time limit was set for participants' responses and the average total time on task was 40 Mn. Eye movements (Tobii 120 hz) and verbal responses were recorded. Two independent raters awarded responses for each item either 1 or 0 points for correct and incorrect answers respectively (interrater agreement = 0.93). Coding criteria were based on the accuracy of each response regarding the type of relation content depicted and on the accompanying explanation given in the textbook. Three categories of Areas of Interest were used per item. (i) RAOI1 for location of relevant information in the left hand graphic, (ii) RAOI2 for location of relevant information in the right hand graphic, (iii) IRAOI for all locations containing irrelevant information. The size of IRAOI was always greater than the size of the two RAOIs.

## Results

Verbal responses scores showed that more than 75% of the paired graphics were understood by participants ( $M = 27.9$ ,  $SD = 4.7$ ). Eye tracking results presented in Figure 2 give a comparison of the total fixation durations on relevant and irrelevant AOIs. An ANOVA for these data revealed a significant effect of AOI type,  $F(2, 34) = 17.70$ ,  $\eta_p^2 = .51$ ,  $p < .0001$ , with relevant AOIs receiving more attention than the irrelevant AOIs,  $F(1, 17) = 96$ ,  $\eta_p^2 = .80$ ,  $p < .0001$ . Participants spent more time on the first picture of the pair but this effect was not significant.

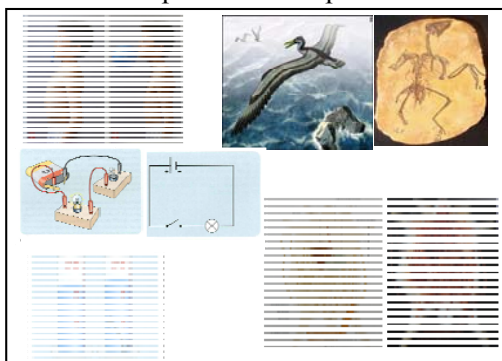


Fig. 1. Examples of paired graphics of the experiment

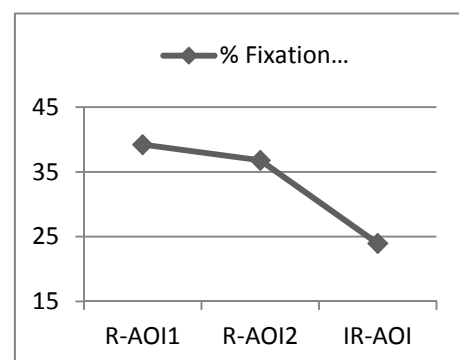


Fig. 2. Mean percent of fixation duration in AOIs

A very high number of transitions between the two relevant AOIs was also recorded. Correlation between time on task and comprehension score was negative, and not significant ( $r(17) = -.31$ ),

showing that time did not really matter. In addition to the quantitative measures reported above, a finer-grained qualitative analysis with respect to individual items was performed by two independent judges. This analysis sought to identify possible features of the paired graphics that could influence how readily they were understood. The variables identified included: (i) *typicality of the item* - classes of paired graphics commonly found in educational materials or particular fields (e.g. realistic versus schematic) were more likely to be understood than less familiar classes (e.g. stages of a specific process); (ii) *number of entities needing to be compared* - comprehension was better when fewer comparisons were required; (iii) *conspicuity of the relation (natural or cued)* – when the aspects to be compared across the two pictures had a relatively high perceptual salience, understanding was more likely (iv) *presence of relevant context* – if the pictures included relevant contextual material restricting the range of possible inferences, comprehension was facilitated; (v) *prior knowledge* – familiar subject matter was better understood than unfamiliar content. However, for these adult participants, neither the complexity of the depicted content nor the number of relations that needed to be determined seemed to have much influence on comprehension performance.

## Conclusion

Unsurprisingly, adult participants in this exploration understood most of the presented paired graphics. They also tended to direct their attention to the high relevance areas within those depictions. A range of variables was identified that appear to influence how readily paired graphics are comprehended. These probably reflect the experience that the participants have built up from repeated exposure to this widely used explanatory visualization format. However, it seems likely that the results for young children who have not yet had this exposure would be somewhat different because of their underdeveloped graphicacy capacities. Given the increasing reliance on graphic forms of information in today's society, this is an issue that should perhaps receive formal attention by educators. It may be that appropriate interventions could be devised to help the acquisition of graphicacy skills, in much the same way that interventions have been used to target literacy and numeracy skills. Studies are currently in progress to determine how young children deal with this same set of materials. Results from the two sets of exploratory studies will be used as the basis for designing more targeted and structured experiments to investigate the nature of graphicacy and how it develops. Regarding paired graphics, it would be helpful to analyse more deeply how learners processed the two graphics (using scan-paths). It would be also very interesting to examine in detail the reasons for the difference in learners processing of the left and right graphics.

## References

- Aldrich, F., & Sheppard, L. (2000). Graphicacy; The fourth 'R'? *Primary Science Review*, 64, 8–11.
- Balchin, W.G.(1985). Graphicacy comes of age, *Teaching Geography*, 11 (1), 8–9.
- Bétrancourt, M., Ainsworth, S., de Vries, E., Boucheix, J.M., & Lowe, R.K. (2012). EARLI SIG, Grenoble, 28-31 August.
- De Vries, E., & Lowe, R. K. (2010, August). Graphicacy: What does the learner bring to a graphic? Paper presented at the EARLI SIG 2 *Comprehension of Text and Graphics meeting, Tübingen, Germany*.
- Postigo, Y., & Pozo, J. I. (2004). On the Road to Graphicacy: The learning of graphical representation systems. *Educational Psychology*, 24(5), 623–644.
- Roth, W.-M., Pozzer-Ardenghi, L., & Han, J. Y. (2005). Critical Graphicacy. *Understanding Visual Representation Practices in School Science Series: Science & Technology Education Library*, Vol. 26. New York: Springer.
- Winn, W.D. (1993). An account for how people search for information in diagrams. *Contemporary Educational Psychology*, 18, 162-185.

### Text-books:

- Spotlight Science 8 (2001) Nelson Thornes; Exploring Science 7 (2008) Pearson Schools
- Science Aspects 1, (2009).G.Linstead, O.Gyder, G. Przywolnik, L. Salfinger, T. Herbert. Pearson Heinemann
- Sciences, Cycle 3; (2003). , JM. Rolando, G. Simonin, P. Pommier, J. Nomblot, JF. Laslaz, S. Combaluzier Magnard, Paris.
- Sciences et technologies. Cycle 3 (2002). Guichard, J., Hebrard, J. (dir.), Zana, B. (2002). Hachette Education, Paris.
- Pulsar, NASK, havo/vwo. (2006). Van den Berg, A., et al. Noordhoff Uitgevers, Groningen, The Netherlands.

# Reading Their Way into Science: Students' Strategies for Comprehending Research Articles

*Isabel Braun, Matthias Nückles*

University of Freiburg, Department of Educational Science

Rempartstraße 11, 79098 Freiburg, Germany

isabel.braun@ezw.uni-freiburg.de, matthias.nueckles@ezw.uni-freiburg.de

**Abstract.** Research articles can inform students about experts' ways of thinking and arguing, thus supporting their development of scientific literacy. However, research articles target an expert audience and students can be expected to experience comprehension difficulties when reading them. The instructional approach of adapting research articles to students' level of understanding is called "adapted primary literature" (APL). Analyses of the strategies students apply when reading research articles (original or APL) are currently missing. Therefore, we collected think-aloud data ( $n=17$ ) and carried out an in-depth analysis of students' strategic processing. Results showed that students applied mostly surface-level and deep-level cognitive strategies but hardly any epistemic reasoning strategies and sourcing strategies. Students who read APL ( $n=8$ ) used significantly higher proportions of epistemic reasoning strategies and deep-level cognitive strategies than students who were given the original research article ( $n=9$ ). We conclude that when students read research articles support should focus on prompting epistemic reasoning strategies and sourcing strategies.

**Keywords:** Reading strategies, learning from science texts, adapted primary literature

## Background and research questions

Research articles serve as primary means of communication among scientists and epitomize the processes of knowledge generation agreed upon by the scientific community (Swales, 1990). Newcomers entering the research arena need to acquire skills for processing and producing them. But even if they do not intend to pursue careers as researchers students could benefit from reading research articles as they can inform them about experts' ways of thinking and arguing (Yarden et al., 2001; Gillen, 2006). As research articles target an expert audience students can be assumed to experience multiple difficulties as they try to comprehend them. Indeed, students seem to struggle when reasoning about the epistemological status of sentences (van Lacuma et al., 2011). And many fail to draw pragmatic inferences (e.g., inferences on authorial intentions) when reading research articles (Haas & Flower, 1988). However, reports on students' reading of research articles contain mostly anecdotal evidence. Thorough analyses of the strategies students apply when reading and trying to comprehend research articles are missing (cf. Goldman & Bisanz, 2002).

There are currently no reports of research articles being used as instructional texts in high school science classrooms. Modified versions of research articles, so called "adapted primary literature" (APL), have been implemented in biology classrooms (e.g., Yarden et al., 2001). APL refers to research articles that retain their rhetorical structure and linguistic features but are modified so as to be comprehensible for high school students (Yarden et al., 2001). For example, specialized terms in the introductory section are replaced or explained whereas details in the method section are eliminated. APL seems a promising approach that could help ease newcomers' enculturation into science reading practices and facilitate the development of scientific literacy. Evidence of what students do while reading APL will allow researchers to capitalize on the benefits of this innovative instructional approach.

In our study, we carried out an in-depth analysis of students' on-line strategies for comprehending research articles, both original and modified. The modified research article was an instance of APL. We investigated the following research questions:

1. Which comprehension strategies do students apply when reading research articles for understanding?
2. How does students' strategic processing of APL differ from their processing of an original research article?

Building on studies of students' reading of popular scientific articles (e.g., Phillips & Norris, 1999) and instructional texts (e.g., DeeLucas & Larkin, 1986) as well as the small set of studies investigating students' reading of research articles (e.g., Haas & Flower, 1988), we expect students to show a high proportion of surface-level cognitive strategies aimed at the propositional content of the texts (re-reading, paraphrasing) and a small proportion of those strategies used by scientists to construct situation models (particularly, critical evaluation against epistemic criteria). We assume that students reading APL show a higher proportion of deep-level cognitive strategies than students reading an original research article. The picture should be reversed for

surface-level cognitive strategies and metacognitive strategies. We derive these assumptions from models of text comprehension and studies that investigated the interactions of textual cohesion and prior knowledge.

## Method

As part of a larger study on learning from science texts, we had 17 high school juniors and seniors (aged  $M=17.47$ ,  $SD=.62$ ) think aloud while reading a research article. The article was published in *Animal cognition* and reports on an experiment investigating tool behaviour of crows. Nine students were given the original research article (2540 words), eight received a modified version (APL, 2517 words). First, the students provided information on their academic achievement. Then they were given instructions on how to think aloud. Then, they were given 35 minutes to study the text assigned to them. The experimenter prompted the students to think aloud if necessary. When they were finished reading, the students answered a questionnaire assessing their comprehension of key ideas and provided ratings of the text.

Before analyzing the think-aloud protocols, we excluded first-time verbatim reproductions of individual sentences or full sections of text because they indicate automatic rather than strategic text processing. We then segmented the think-aloud protocols into idea units. Every segment was assigned to one of the categories illustrated in Table 1.

Table 1

Taxonomy of comprehension strategies for analysis of think-aloud protocols

Table 2

Students' use of strategies (relative frequencies)

Category	Description	M (SD)	
Epistemic reasoning	Evaluating statements against epistemic criteria (e.g., "Well, sounds like the researchers are pretty sure of them [the findings] since they say in the last paragraph: Well, here we were able to show that crows undoubtedly [...]"), abstracting epistemological assumptions underlying text (e.g., "Well, I believe studies like this can always change.")	2.12%	(2.53)
Sourcing	Predicting author or publication, abstracting function of text, inferring authorial intention (e.g., "I would say a very optimistic portrayal of the – a bit of an advertisement for behavioural research.")	0.16%	(0.45)
Surface-level cognitive	Re-reading, paraphrasing	24.7%	(16.41)
Deep-level cognitive	Establishing connection between ideas in text, establishing connection to prior knowledge, forecasting, etc.	25.15%	(11.95)
Meta-cognitive	Monitoring comprehension, exploring comprehension failure, regulating reading in face of comprehension failure	12.31%	(8.97)
Impasse	Voicing breakdown in comprehension	12.34%	(10.2)
Interjection	Stand-alone utterance of obscure intent (e.g., "yup", "oh my", "well")	18.82%	(9.6)

## Results

On average, the students voiced 79.47 ( $SD=37.52$ ) thoughts while reading the original research article or APL. The total number of thoughts did not differ significantly between the two conditions. A repeated-measures ANCOVA (with academic achievement in native-language reading and writing instruction as covariate) revealed significant differences in relative frequencies between the categories of comprehension strategies (Table 1),  $F(90,6)=4.59$ ,  $p<.001$ ,  $\eta^2=.23$ . Surface-level and deep-level cognitive strategies and interjections dominated students' thoughts as they tried to comprehend the research article (Table 2). Very small proportions of the thoughts voiced by the students indicated epistemic reasoning strategies and sourcing strategies (Table 2).

We tested for differences in strategic processing between the two conditions using MANCOVA (with academic achievement in native-language reading and writing instruction as covariate). The analysis showed that the students who read APL employed epistemic reasoning strategies and deep-level cognitive strategies to a significantly higher extent than the students who were given the original research article (Figure 1),  $F(14,2)=3.76$ ,  $p<.05$ ,  $\eta^2=.35$  and  $F(14,2)=8.59$ ,  $p<.005$ ,  $\eta^2=.55$ .

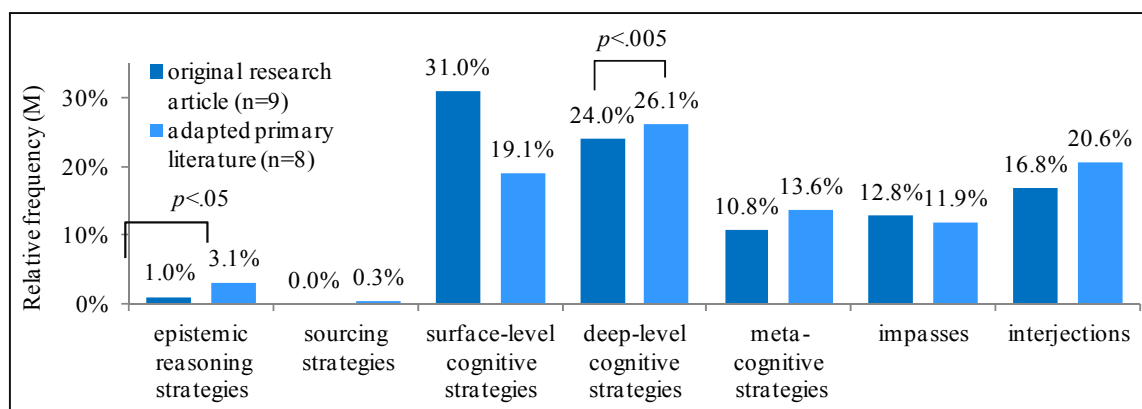


Figure 1. Students' strategies for comprehending research articles by condition

## Conclusions

The results of our study support our assumption that students focus on propositional content when reading and trying to comprehend research articles: The students allotted a substantial portion of their engagement with the texts to re-reading and paraphrasing. Higher-level engagement with the texts was largely restricted to strategies such as establishing connections and forecasting, with very little indication of students' using epistemic reasoning strategies and sourcing strategies. In addition, a substantial portion of the thoughts students' voiced while reading, particularly the interjections, did not appear strategic and directed at comprehension. In summary, our findings suggest that students draw on the strategies they display when reading and trying to comprehend popular scientific articles and instructional texts (e.g., DeeLucas & Larkin, 1986; Phillips & Norris, 1999). They might, therefore, miss opportunities for learning provided by research articles, that is, opportunities for learning about the inner workings of science – thought styles, discourse rules and the role of research articles in generating and establishing knowledge claims.

As we had hypothesized, APL facilitated higher-level engagement with the research article, probably by tying fewer resources for establishing local coherence than the original research article. However, despite showing significantly higher proportions of deep-level cognitive strategies than the students who read the original research article, the students who were given APL remained low-frequency users of epistemic reasoning strategies. Discrepant to our assumption, the proportion of metacognitive strategies was not higher with the original research article than with APL. Based on the high proportion of re-reading and paraphrasing we observed for the original research article, we assume that the students used these strategies to regulate their reading when they experienced breakdowns in comprehension. The execution of these strategies might be automatized, explaining the lower proportion of explicit metacognitive processing.

In conclusion, our findings indicate that when given a research article students do a lot of what they are trained to do in school in order to learn from science texts but little of what scientists do when reading scholarly scientific texts. APL seems to ease students' reading their way into science but prompting of epistemic reasoning strategies seems necessary and should be explored in future studies.

## References

- Bazerman, C. (1988). *Shaping written knowledge*. Madison, WI: University of Wisconsin Press.
- DeeLucas, D., & Larkin, J. H. (1986). Novice strategies for processing scientific text. *Discourse Processes*, 9, 329-354.
- Gillen, C. M. (2006). Criticism and interpretation: Teaching the persuasive aspects of research articles. *CBE-Life Sciences Education*, 5, 34-38.
- Goldman, S. R., & Bisanz, G. L. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León & A. C. Graesser (eds.), *The psychology of science text comprehension* (pp. 19–50). Mahwah, NJ: Erlbaum.
- Haas, C., & Flower, L. (1988). Rhetorical reading strategies and the construction of meaning. *College Composition and Communication*, 39, 167-183.
- Phillips, L. M., & Norris, S. P. (1999). Interpreting popular reports of science: What happens when the reader's world meets the world on paper? *International Journal of Science Education*, 21(3), 317-327.
- Yarden, A., Brill, G., & Falk, H. (2001). Primary literature as a basis for a high-school biology curriculum. *Journal of Biological Education*, 35(4), 190-195.



# Limits and Potentials of Bottomup Processing of Multiple External Chemical Representations

*David Corradi*

University of Leuven (KU Leuven)  
Dekenstraat 2 bus 3773, 3000 Leuven, Belgium  
david.corradi@ppw.kuleuven.be

*Jan Elen*

University of Leuven (KU Leuven)  
Dekenstraat 2 bus 3773, 3000 Leuven, Belgium  
jan.elen@ppw.kuleuven.be

*Geraldine Clarebout*

University of Leuven (KU Leuven)  
Dekenstraat 2 bus 3773, 3000 Leuven, Belgium  
geraldine.clarebout@ppw.kuleuven.be

**Abstract.** When learning with Multiple External Representations (MER), low prior knowledge learners are said to be limited in using these MER to achieve conceptual understanding. Yet little is known about what these limitations precisely entail. To understand this, we presented 101 low prior knowledge level learners with chemical MER to see a) how many, and what kind of ideas (propositions) learners remembered from these MER and b) how these propositions impact conceptual understanding of the content. ANOVA and regression analysis found that it is not the amount of propositions; more important to achieve conceptual understanding is whether those propositions are correct or not, and whether these are propositions that connect the MER to each other.

**Keywords:** Chemical Representations, Prior knowledge, Conceptual Understanding, Propositions

## Introduction

Research has claimed that integrating different kinds of information from diverse representational formats (Multiple External Representations (MER), i.e., text, symbols and images) into a coherent structure is necessary to fully understand the concepts and principles of a scientific domain (Berthold & Renkl, 2009). Models on learning with representations have suggested that the combination of descriptive representations (e.g., texts, symbols) and depictive representations (e.g., images) can be beneficial for learning (Carney & Levin, 2003; Schnotz & Bannert, 2003). According to the model of Schnotz on text-picture comprehension (Schnotz, 2005), this is because linguistic propositional representations can interact with non-linguistic mental models to form coherent schemata and achieve conceptual understanding (Seufert, 2003).

Previous research (e.g., Corradi, Elen & Clarebout, 2012) with low prior knowledge learners has found limitations when these external representations are scientific representations. Explanations focus on limited schema of low prior knowledge learners and limited attention towards other representations than the easiest accessible representations (e.g. a text; Schnotz & Bannert, 2003). Yet, having low prior knowledge with limited schemata on the subject or giving little attention towards a representation does not necessarily mean that the information is not processed. These explanations, in other words, remain at the superficial level and do not detail, for example in terms of the model of Schnotz, why precisely learners lack the ability to use MER to increase their understanding. Secondly, based on these explanations, we have little understanding on how to enhance the use of MER to increase deep understanding with low prior knowledge learners.

We know from previous studies on scientific (in our case – chemical) representations that learners filter and (mis)interpret information in a specific way (e.g., Treagust, Chittleborough & Mamiala, 2002). This study builds further on these studies by investigating the precise effects of scientific (M)ER on the creation of the internal representation, with low prior knowledge learners in the domain chemistry. MER in our study are texts (as clarifying representations), symbolic representations and

submicroscopic representations. We want to see whether these MER give an advantage for processing and understanding the information compared to only text. We also want to understand how the internal representation influences deep conceptual understanding of the represented information. Our research questions are: 1) Is there an advantage of the combination of depictive and descriptive representations for both increase of understanding and the type and amount of information that is being filtered? and 2) how does the internal representation information influences gains in conceptual understanding? To understand how learners filter ideas from representations, we will use the notion of propositions, which we operationalize as segments of a representation involving a single idea (Recht & Leslie, 1988). To understand what the impact is of each representation(al combination) we predict that the group that receives texts with symbolic and submicroscopic representations will differ significantly concerning learning gains with the groups that receive only two representations (text+symbolic; text+submicroscopic) or only one (text-only) (hypothesis 1). Concerning the relationship between the propositions and the construction of the mental model we predict that amount of propositions will influence the amount of correct propositions (hypothesis 2); and we predict that learners will have more propositions or more correct propositions deduced from a single representation compared to learners who receive the information through MER (hypothesis 3). Finally we hypothesize that the amount of correctly remembered propositions influences the increase in conceptual understanding (hypothesis 4).

### *Methods*

*Participants and design.* Participants were 101 first and second year college students in Socio-educational Care Work. The study is a pre-post randomized experiment. We focused on chemistry, and our MER were texts, with 1 or 2 additional representations, namely symbols and submicroscopic representations (see, Corradi, et al., 2012). Group 1 (n=23) was the text-only condition (control group). Group 2 (n=25) received texts and symbolic representations. Group 3 (n=26) received texts and submicroscopic representations. Group 4 (n=27) received texts with symbols and submicroscopic representations.

*Materials, Instruments and Procedure.* Pre and post-tests measured understanding of basic chemical concepts. Materials were a 1300-word text and, depending on the condition, learners received texts either alone or with one or two types of representations added, whereby each representation was printed on two pages. After studying the MER, learners received a short intermediate task, then the assignment to write down in short sentences (no drawings or symbols) what they remembered from the intervention, before they received the post-test.

All information written down in the assignment after the intermediate task, were recoded into propositions (defined as a single idea expressed through minimum a verb and a noun; Recht & Leslie, 1988), which can be counted (amount of propositions), analysed whether correct or wrong (correct propositions), what representation they originated from and whether it originated from one or more representations. This allowed us to identify the precise effect of single versus multiple external representations on learners with limited schemata of chemistry. Additionally, this allowed us to observe the relationship between mental model construction and conceptual understanding.

### *Results and Discussion*

Groups did not differ on the pre-test results,  $F(3,97) = .269$ ,  $p = .848$ ,  $\eta^2_p = .008$ . Most learners increased their understanding of chemistry significantly,  $t(100) = 12.601$ ,  $p < .001$ . However, there was no significant difference between the different groups, concerning the learning gains (difference between pre- and post-test),  $F(3,97) = 1.552$ ,  $p = .206$ ,  $\eta^2_p = .046$  (hypothesis 1 rejected). There was a significant difference in the amount of propositions (hypothesis 2 confirmed),  $F(3,97) = 3.662$ ,  $p = .016$ ,  $\eta^2_p = .101$ .

TukeyHSD post hoc found that the group who received information through texts and submicroscopic representations scored significantly higher than groups who received all three representations (6.39,  $p = .016$ ). There was also a marginal significant difference between text-only (group 1) and a marginal non-significant difference text with submicroscopic representations (group 3) (5.58,  $p = .060$ ). We found a marginal difference for the correct propositions,  $F(3,97) = 2.633$ ,  $p = .054$ ,  $\eta^2_p = .075$ . However, seen in proportion (correct propositions divided by all propositions), no significant difference was found between the groups (hypothesis 3 rejected),  $F(3,97) = 1.618$ ,  $p = .19$ ,  $\eta^2_p = .048$ . We also found that this variable (correct propositions divided by all propositions) is the only predictor of increase in conceptual understanding (.30) with  $R^2 = .09$ ,  $F(1, 100) = 7.44$ ,  $p = .008$  (hypothesis 4 confirmed). Next to this, learning gains of conceptual understanding correlated with the amount of overlapping propositions remembered, .547,  $p = .003$ , which means that it is not the amount of ideas remembered, but their specific nature (whether correct or not and whether overlapping or not) that provides indications of how well learners formed a coherent schemata of the information. In the light of the model of Schnotz and Bannert (e.g., 2003) we found an advantage of the combination of descriptive plus depictive compared to other groups concerning the amount of ideas remembered, this did not give an advantage for their understanding of chemistry. Forming schemata of abstract chemical concepts cannot be solely based on combining depictive and descriptive representations, but learners in the group who received a text combined with a submicroscopic representation do seem to have an advantage over the other groups as they remember significantly more. Learners in the group with three representations remembered the least amount of ideas, but the more overlapping propositions were remembered (overlapping between the three representations) the more they increased their understanding of chemical concepts. Our most important conclusion is that low prior knowledge learners are sensitive towards learning with multiple representations, and our results indicate that they do hold the potential to achieve conceptual understanding.

## References

- Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, 101(1), 70–87.
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14(1), 5–26.
- Corradi, D., Elen, J., & Clarebout, G. (2012). Understanding and enhancing the use of multiple external representations in chemistry education. *Journal of Science Education and Technology*. Online First. doi: 10.1007/s10956-012-9366-z
- Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 60(1), 16–20.
- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13, 227–237.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed), *The Cambridge Handbook Of Multimedia Learning*, (pp. 49-69). New York, NY: Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13, 141–156.
- Treagust, D. F., Chittleborough, G. D., & Mamiala, L. T. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, 24, 357–368.

# Automated Analysis of Pupils' Self-Explanations of a Narrative Text

Philippe Dessus<sup>1</sup>, Maryse Bianco<sup>1</sup>, Aurélie Nardy<sup>1</sup>, Françoise Toffa<sup>1</sup>, Mihai Dascalu<sup>2</sup>,  
Stefan Trausan-Matu<sup>2</sup>

<sup>1</sup>LSE & IUFM, Grenoble-1 and Grenoble-2 Universities, France

{Philippe.Dessus, Maryse.Bianco, Aurelie.Nardy}@upmf-grenoble.fr, Francoise.Toffa@gmail.com

<sup>2</sup>“POLITEHNICA” University of Bucharest, Romania

{Mihai.Dascalu, trausan}@cs.pub.ro

**Abstract.** We present a method devised to automatically analyze pupils' verbalizations during reading and to reveal some of the strategies they use. An experiment with 44 primary school pupils (3<sup>rd</sup> and 5<sup>th</sup> grade) reading a narrative text and verbalizing what they understood at predefined breaks showed, firstly, that machine results are correlated with experts'; secondly, that the recall of recent sentences is not uniform across verbalizations and subject to a grade effect; thirdly, that there is a grade effect in the recall of distal causal sentences.

**Keywords:** Self-Explanations, Reading Strategies, Latent Semantic Analysis, Comprehension

## Introduction

The use of strategies during reading is widely recognized as a crucial determinant of reading comprehension. Second degree and high school pupils who are good comprehenders are mostly strategic readers (Graesser, 2007). These strategies can be elicited through self-explanations and have been categorized by McNamara (2004) as follows: comprehension monitoring, paraphrasing, elaboration, prediction, and bridging. One important skill these strategies exploit is to be able to establish semantic and causal relationships between the read sentences (Wolfe, Magliano, & Larsen, 2005).

Based on these findings, McNamara et al. (2007) developed *iSTART*, a cognitive tutor that automatically categorizes self-explanations, partly using Latent Semantic Analysis (Landauer & Dumais, 1997). Any thorough analysis of self-explanation reports is a very demanding and subjectivity-oriented activity, and the use of systems like *iSTART* to detect pupils' reading strategies is more than challenging. Since a cognitive tutor guides the reader through pre-defined steps alternating between reading and verbalizations, we consider its use as a scenarization of the reading and comprehension process. This computer-based scenarization is made possible through the wide range of reading strategies and the feedback possibilities (Vitale & Romance, 2007). At least two kinds of reading scenarios are considered: *pedagogical scenarios*, whose aim is to orchestrate the reading activities, and *assessment scenarios*, targeted at supporting the way learners' reading is assessed.

The aim of this paper is to focus on the second kind of scenarios and to present preliminary results for the basis of a cognitive reading tutor. We introduce and test a way to automatically analyse pupils' self-explanations of what they have read, through an LSA-based analysis of word usage. Since paraphrases are one of the most frequently used strategies (McNamara, 2004), we focus our study on how two main kinds of sentences are paraphrased: focal (the latest sentence before a verbalization) and causal sentences (identified by a causal analysis of the text), because it is worth distinguishing the mere paraphrase of the latest read sentence and more elaborated paraphrases, involving a deeper comprehension of the read text. This paper investigates novel research paths. Firstly, it focuses on elementary school pupils, a category of students seldom investigated. Secondly, an LSA space is chosen that fits best the pupil's knowledge by using a corpus composed by Denhière et al. (2007), comprising 3.2 million words and validated by a test involving association norms. Thirdly, we propose a dynamic view of the self-explanation process, in analysing the verbalizations at different break points throughout the story.

Our research questions are firstly to compare human expert categorization of paraphrases to the semantic similarity between text sentences and self-explanations, obtained by means of LSA.

Secondly, we expect a “recency effect”, stating that the information children self-explain most often pertains to very close sentences to the verbalization break. Thirdly, we will investigate the way pupils account for causal relations (either local or distal) in retelling causally-related text sentences.

## Method

### *Participants*

22 third and 22 fifth grade pupils from the same school and from a middle socio-economic background participated in our experiment.

### *Materials and Procedure*

One narrative text was read and self-explained by the pupils: *Matilda* (453 words, 6 self-explanation breaks). The text was chosen to be within the reading level of participants, so that differences in verbalizations would indicate differences in reading strategies instead of comprehension difficulties. In order to perform a fine-grained analysis, the initial text was split in 45 segments (of about 1 sentence each). We performed a causal analysis so that both *local* (when the causal antecedent is close to the reference sentence) and *distal* antecedents (when the causal antecedent is somewhat farther, out of the reader’s working memory) of sentences were determined as in Millis et al. (2006). We finally performed a propositional analysis of the text, which allowed us to extract macro-propositions and to support the coding of what was remembered by the participants.

Participants individually read the text out loud and stopped at predetermined breaks to self-explain the text segment just read, the whole activity being recorded. The task was explained to pupils as follows: “During your reading you will stop at each icon to tell out loud what you have understood, just at this time”. Their verbalizations were then transcribed and each self-explanation was semantically compared with LSA—all the text sentences before the self-explanation breaks. Two of the co-authors analysed pupils’ verbalizations proposition by proposition and categorized them according to McNamara’s (2004) coding scheme. Disagreements were also discussed and resolved by consensus.

## Results

First of all, we computed accuracy measures in order to compare human vs. LSA values of sentence relatedness and to check the validity of the computer-based measures. Pearson correlations between the number of paraphrases per verbalization ( $V_n$ ) detected by the two raters and LSA similarities between each verbalization and the previous sentences were as follows:  $V_1$ :  $r=.48$ ;  $V_2$ :  $r=.58$ ;  $V_3$ :  $r=.74$ ;  $V_4$ :  $r=.29$ ;  $V_5$ :  $r=.57$ ;  $V_6$ :  $r=.61$ , which shows that human judgments of paraphrases expressed by children on each paragraph are moderately to strongly related to LSA measures of similarities. We then investigated the extent to which each self-explanation was related to the last read sentence (focal). Figure 1a presents the cosine similarities, processed by LSA, between each verbalization and the focal sentence, by grade. We observe that the recency effect varies across verbalization plots, indicating that this effect is dependent of the content conveyed by the last sentences. Moreover, the focal sentence, in general, does not have a higher similarity with the related verbalization than the average of other previous sentences, except for  $V_4$ :  $t(43)=7.5$ ,  $p < .0005$ . Two-way ANOVAs showed a significant difference between grades for  $V_6$ ,  $F(1, 42)=7.01$ ;  $p < .05$  and a tendency for  $V_2$ ,  $F(1, 42)=3.22$ ,  $p < .09$ . Although grade 3 pupils tended to recall the last sentence at these points more frequently, the semantic content of the last sentence seems to be the main determinant of focal recall.

The third hypothesis predicted that the semantic content of local and distal sentences, as determined by the causal analysis, is more often verbalized than the rest of the previous text and the focal sentence. Moreover, the local-centered causal sentences were expected to be better recalled than the distal-centered ones (see Figure 1b, depicting the LSA-based similarities between  $V_n$  and their

related causal sentences). Results first showed that local and distal causal sentences are, in all cases but two (local vs.  $V_1$  and  $V_5$ ), significantly more verbalized than the rest of the text. Moreover, the content of *local* causal sentences was significantly better recalled than *focal* sentences in  $V_1$  and  $V_3$  (resp.  $t(43)=3.11$ ,  $p < .005$ ;  $t(43)=9.45$ ,  $p < .0005$ ). Unexpectedly, the content of *distal* causal sentences was better recalled than *local* causal sentences for  $V_1$ :  $t(43)=6.09$ ,  $p < .0005$ ;  $V_2$ :  $t(43)=8.49$ ,  $p < .0005$ . Two-way ANOVAs showed significant differences between grades for  $V_1$  (distal),  $F(1, 42)=4.43$ ,  $p < .05$ ; and a tendency for  $V_6$  (distal),  $F(1, 42)=3.90$ ,  $p < .06$  and for  $V_3$  (local),  $F(1, 42)=2.91$ ;  $p < .1$ . Overall, participants' strategies focused on causality, rather than recency.

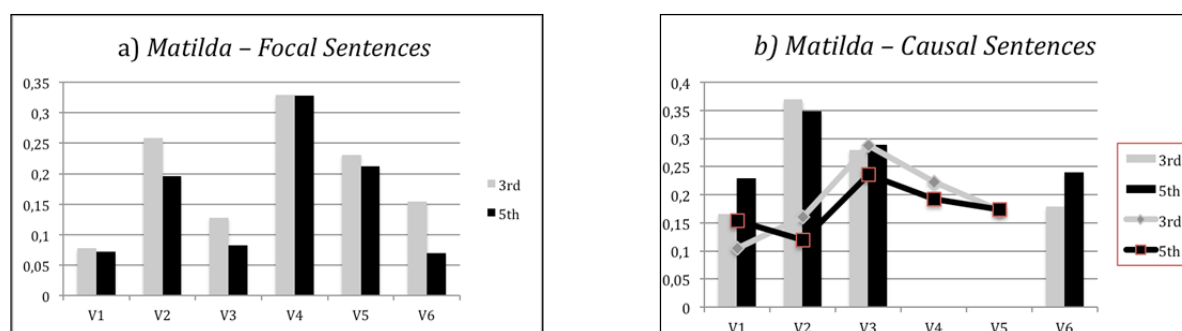


Figure 1 a) Mean LSA-based values for similarity of focal sentences by grade; b) Mean LSA-based values for similarity of causal sentences, by grade. Lines: local causality; bars: distal causality.

## Discussion

This study presented a first attempt to set up the foundations of a cognitive reading tutor aiming at analysing pupils verbalizations to get some traces of their strategies. Results showed that LSA-based analyses of verbalizations correlate moderately to high with those of human experts. Additionally, and as also shown by Trabasso and van den Broek (1985), participants tended to recall sentences they read according to causality-driven, rather than recency-driven strategies, which reveal to some extent their comprehension strategies. Eventually, there was a grade effect on the way distal and local causal sentences are recalled. Further research will aim to refine and to improve the validity of automated analyses, as well as to combine the verbalization-based with complexity-based ones, in order to calibrate a pupil's reading performance.

## References

- Denhière, G., Lemaire, B., Bellissens, C., & Jhean-Larose, S. (2007). A semantic space for modeling children's semantic memory. In T. K. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 143–165). Mahwah: Erlbaum.
- Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies* (pp. 3–26). Mahwah: Erlbaum.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem. *Psychol. Rev.*, 104(2), 211–240.
- McNamara, D. (2004). SERT: Self-Explanation Reading Training. *Discourse Processes*, 38, 1–30.
- McNamara, D., Boonthum, C., & Levinstein, I. (2007). Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. In T. K. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 227–241). Mahwah: Erlbaum.
- Millis, K., Magliano, J., & Todaro, S. (2006). Measuring discourse-level processes with verbal protocols and Latent Semantic Analysis. *Scientific Studies of Reading*, 10(3), 225–240.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612–630.
- Vitale, M. R., & Romance, N. R. (2007). A knowledge-based framework for unifying content-area reading comprehension and reading comprehension strategies. In D. S. McNamara (Ed.), *Reading comprehension strategies* (pp. 73–104). Mahwah: Erlbaum.
- Wolfe, M. B. W., Magliano, J. P., & Larsen, B. (2005). Causal and semantic relatedness in discourse understanding and representation. *Discourse Processes*, 39(2-3), 165–187.

# Students' Eye Movements when Solving Mathematical Word Problems together with Illustrations

*Tinne Dewolf*

Centre for Instructional Psychology and Technology, KU Leuven  
Dekenstraat 2, box 3773, 3000 Leuven, Belgium  
Tinne.Dewolf@ppw.kuleuven.be

*Wim Van Dooren*

Centre for Instructional Psychology and Technology, KU Leuven  
Dekenstraat 2, box 3773, 3000 Leuven, Belgium  
Wim.VanDooren@ppw.kuleuven.be

*Frouke Hermens*

Centre for Experimental Psychology, KU Leuven  
Tiensestraat 102, box 3711, 3000 Leuven, Belgium  
hermensf@gmail.com

*Lieven Verschaffel*

Centre for Instructional Psychology and Technology, KU Leuven  
Dekenstraat 2, box 3773, 3000 Leuven, Belgium  
Lieven.Verschaffel@ppw.kuleuven.be

**Abstract.** In the present study we measured students' eye movements to further investigate – as observed by Dewolf, Van Dooren, EV Cimen and Verschaffel (2011) – why representational illustrations do not help pupils to respond more realistically to problematic word problems (P-items), i.e., problems that require taking the realities of the context into account. By measuring eye movements we hoped to gain insight into how P-items are solved, and if attention is being paid at the illustrations. Thirty higher education students solved 16 word problems while their eye movements were measured. They were equally divided in three conditions; the problems together with representational illustrations (illustrations that represent the problematic situation), with decorative illustrations (illustrations that have no link with the problem) and without illustrations. We found that also higher education students neglect their everyday knowledge when solving P-items, and that they barely look at the illustrations next to the problems.

**Keywords:** Word problem solving, representational illustrations, eye movement research, mathematics class

## Introduction

In previous research, it has repeatedly been shown that elementary school pupils tend to exclude realistic considerations when solving problematic word problems (P-items), i.e., problems that, in contrast to classical standard word problems (S-items), require taking the realities of the situation described in the problem into account (Verschaffel, De Corte, & Lasure, 1994; Verschaffel, Greer, & De Corte, 2000). An example of such a P-item is: “A man wants to have a rope long enough to stretch between two poles 12m apart, but he has only pieces of rope 1.5m long. How many of these pieces would he need to tie together to stretch between the poles?” (Verschaffel et al., 1994). During the past 15 years several researchers have tried various manipulations and interventions to better understand the origin of this tendency and to increase the number of realistic reactions (RRs) to P-items. Dewolf, Van Dooren, EV Cimen and Verschaffel (2012), for example, investigated whether the provision of illustrations that represent the problem situation and/or of a warning that alerts about the tricky nature of some of the problems, increased the number of RRs on P-items. Neither the illustration nor the warning, or even the combination of both manipulations, had a positive impact.

The present study further investigated why providing P-items with an illustration that aims at representing the problem situation and at evoking the solver's real-world knowledge about that situation (a representational illustration according to the categorisation of Carney & Levin, 2002), does not result in more RRs to these P-items. Apart from collecting students' responses to a set of P-items and S-items, we registered their reaction times (RTs), eye movements on the text and the

illustration, and confidence scores. By registering students' eye movements, we hoped to gain more insight in how P-items are understood and solved, and what role the illustrations play. By means of these data we investigated the following three possible explanations for the absence of a positive effect of the representational illustrations, raised by Dewolf et al. (2012): 1) students do not look at the illustrations, 2) students do look at the illustrations but do not experience the problematic character of the P-items, or 3) students do look at the illustrations and build a realistic situational model but nevertheless prefer a non-realistic reaction (NR) due to their beliefs about solving word problems in the mathematics class.

## Method

Thirty higher education students were individually confronted with eight P-items and eight parallel S-items from the study of Verschaffel et al. (1994). The students were equally divided in three conditions. In the Representational Illustration condition (RI-condition), students received the word problems together with representational illustrations, i.e., illustrations that represent the situation described in the word problem. In the No Illustration condition (NI-condition), the word problems were presented in their original form without illustrations. In the Decorative Illustration condition (DI-condition), students received the word problems together with decorative illustrations, i.e., illustrations that have no link with the content of the problem, but just decorate the page.

Students were seated in front of a computer that was attached to the Eyelink II, that measured their eye movements. The word problems, with or without an illustration (depending on the condition), appeared one by one on the computer screen, separated by a fixation point in the middle of the screen. Students had to respond to each problem by verbally stating their answer and possible additional comments, and then had to press a button to go to the next problem. The word problem was presented on the left side of the screen (text area), and the illustration – representational, decorative, or a blank space, depending on the condition – was presented on the right side of the screen (illustration area). After solving the problems, the eye movement device was turned off and the student received a paper-and-pencil questionnaire with the same 16 word problems, presented again with a representational or decorative illustration or without an illustration. In this questionnaire students were asked to indicate, for each item, to what extent they had hesitated about their answer (by responding: a) not at all, b) almost not, c) a little bit, and d) very much), and if so, why.

## Results

First, we found that, just like elementary school pupils, higher education students tend to exclude real world knowledge when solving P-items. Only 27.9% of the reactions to the P-items were realistic. Like in the study of Dewolf et al. (2012), no effect of representational illustrations was found; the number of RRs in the RI-condition (31.3%) did not differ significantly from the number of RRs in the DI-condition (31.3%) or the NI-condition (21.3%) ( $X^2(2,240)=4.38$ ,  $p=.112$ ). Second, there were no significant differences between conditions in RT ( $X^2(2,480)=0.29$ ,  $p=.865$ ). There was however a significant difference between solving P-items and S-items; P-items were solved significantly slower than S-items ( $X^2(1,480)=40.42$ ,  $p<.001$ ). There was no interaction between condition and item type ( $X^2(2,480)=0.77$ ,  $p=.680$ ). So, P-items were processed differently than S-items, c.q. more deeply, irrespectively of the presence or nature of illustrations. Third, students' eye movements on the text showed no significant effect of condition on the number of fixations ( $X^2(2,480)=1.36$ ,  $p=.507$ ). However, significantly more fixations on the text were needed to solve P-items than S-items ( $X^2(1,480)=26.23$ ,  $p<.001$ ), while there was no interaction between condition and item type ( $X^2(2,480)=0.62$ ,  $p=.732$ ). Concerning mean duration of the fixations, there was no effect of condition ( $X^2(2,19207)=0.99$ ,  $p=.610$ ), item type ( $X^2(1,19207)=2.80$ ,  $p=.094$ ), nor an interaction effect ( $X^2(2,19207)=0.77$ ,  $p=.681$ ). Fourth, students' eye movements on the illustration area revealed that



they barely looked at the illustrations. There were however significantly more fixations on the illustration area in the RI-condition (1.5% of all fixations), than in the DI-condition (0.4%), where the illustrations were not linked to the problem, or the NI-condition (0.0%), where the illustration area was blank ( $X^2(1,13230)=34.48$ ,  $p<.001$ ). To calculate how many illustrations in total were at least minimally processed (<50ms), the longest fixation on each illustration, for each item per student was identified. In the RI-condition 28.1% (17.5% P-items and 10.6% S-items) of the illustrations were processed, versus only 11.9% (7.5% P-items and 4.4% S-items) in the DI-condition. This difference between conditions was significant ( $p=.027$ ), leading to the conclusion that more illustrations in the RI-condition were processed than in the DI-condition. We also looked at the relation between looking at the illustration and giving a RR. In both the RI-condition and DI-condition there was no significant relation. Fifth, students' responses on the questionnaire showed that hesitations about the correctness of the answer occurred significantly more on P-items than on S-items, but the amount of hesitation did not differ significantly between the conditions. When looking at the P-items that were solved non-realistically, we see that students tend to hesitate more about the problematic nature of the P-items in the RI-condition than in the other two conditions.

### Conclusion and discussion

The combined analysis of students' responses, RTs, eye movements and confidence scores on a set of P-items and their parallel S-items, leads to the conclusion that representational illustrations of P-items do not help, above all, because students do not look at the illustrations. Two additional explanations are that students did look at them but without noticing the realistic modelling complexity, or that students' beliefs about solving school word problems prevented them from giving a RR even though they did look at the illustration and did notice the realistic modelling complexity in the item.

### References

- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5-26.
- Dewolf, T., Van Dooren, W., EV Cimen, E., & Verschaffel, L. (2012). *The impact of illustrations and warnings on solving mathematical word problems realistically*. Manuscript submitted for publication.
- Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, 4, 273-294.
- Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse, The Netherlands: Swets & Zeitlinger.

## De-symbolization in Learning a New Symbolic System

Gilles Dieumegard

LIRDEF - IUFM

62, r. Vincent Faïta – F 30 000 Nîmes - France

gilles.dieumegard@montpellier.iufm.fr

**Abstract.** We have studied the course-of-action of subjects who had to learn conventional relations in a new system of symbolic representations (Maya numbers). The subjects alternated between processing symbolic operations and intuitive operations with these representations, treating it sometimes rather as signs conventionally representing numbers, and sometimes rather as successions of drawings with visual features – they de-symbolize these representations.

**Keywords:** representation, sign, symbol, conventional relations, visual features

## Introduction

Following Peirce, symbols are signs with an arbitrary structure and conventionally related to the content they represent. But any convention has to be learned, and one of the purposes of schooling is precisely to transmit the conventional relations between symbolic representations and objects or events for which they stand. So we can ask how students use external representations when they have to learn new symbols, *i.e.* to develop a use of new conventional relations between signs and objects.

## Method

We have studied how subjects solve a problem in which they have to learn a new symbolic system: Maya numbers system. It was chosen because, despite its simplicity, subjects of different ages could not only apply previously learned routines, independently of their former mathematical knowledge.

Mayas leave before in Central America. They employ only three symbols in order to write numbers. Here's the Maya writing for numbers 1 to 25. Write in the Mayas' way the continuation of numbers until 50. How do the Mayas write number 100? And 120? 137? 169? 200?

Figure 1. Mayas' problem and the given numbers table

We used a method from cognitive ergonomics which allowed us to analyse the progress of subjects' activity: the course-of-action method (Theureau, 2003). While working on the problem, subjects were asked to think aloud and were videotaped. Then we identified units of action relying on subjects' words, writings and gestural behaviours; their succession drew the subjects' course-of-actions. For this study, we distinguished two types of units:

- symbolic operations, when the subjects interpreted the signs (from the given table or those they wrote) relying on identified relations between elementary signs (*e.g.*, a line) and numbers, or at least considered composed signs as graphic ensembles related to identified numbers;
- intuitive operations, when the subjects imagined or wrote signs (elementary or composed) without knowing the numbers they represented. We observed two types of these operations: a) the subjects increased one-by-one the quantity of some elementary signs; b) the subjects copied entirely or partly composed signs. In order to not overestimate these intuitive operations, we categorized so some units only if we had a proof that subjects, at this moment, didn't relate signs (elementary or composed) with identified numbers.

In this paper, we present an analysis for two subjects: Jonathan, aged 15, and Mathilde, aged 9.

## Results

Both Jonathan's and Mathilde's courses-of-action could be divided in three phases: 1) exploring Maya's signs and numbers table; 2) writing numbers from 26 to 50; 3) writing separate numbers (100, 137, 169, 200). But if Jonathan achieved the given problem within 12 minutes, Mathilde produced only a partially right answer after 24 minutes. Differences in the intertwining of symbolic and intuitive operations in the different phases gave explanations for their different results.

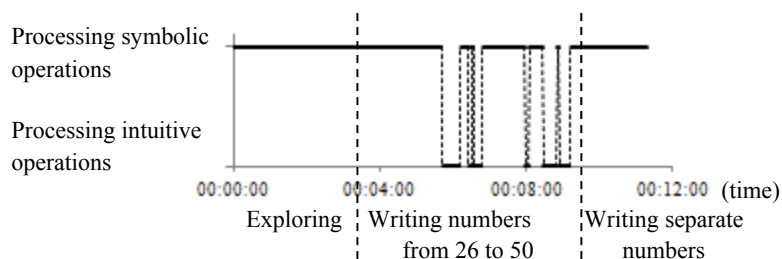


Figure 2. Symbolic and intuitive operations in Jonathan's course-of-action

During the exploring phase (0:00 to 3:31) and the beginning of the writing numbers phase (3:31 to 5:41), Jonathan processed only symbolic operations: first he tried to interpret the given text and Maya numbers table, and second he prefigured the writing of numbers 26 to 50. So he stabilized symbolic relations between Maya signs and numbers: *"That is – that [points the dot and then the shell in the fifth compartment of the fourth line of the table]<sup>2</sup> – it makes one time twenty"* (4:34).

In the second part of the writing numbers phase (5:41 to 9:11), intuitive operations, independent to the symbolic role of the written signs, could be observed. : *"One line -two dots [finishes Maya number 27]; line - one, two, three dots [finishes Maya number 28]; line - one, two, three, four dots [finishes Maya number 29]; line - two lines [finishes Maya number 30]<sup>3</sup>"* (5:41).

These temporary intuitive processes allowed Jonathan to write quickly (1min 43sec) 19 of the 25 numbers from 26 to 50. Nevertheless, from time to time he shifted back to symbolic processes in order to check his writing, in particular to ensure to write in a right way the number 40. Finally he wrote separate numbers (100, 137, 169, and 200) by processing only symbolic operations.

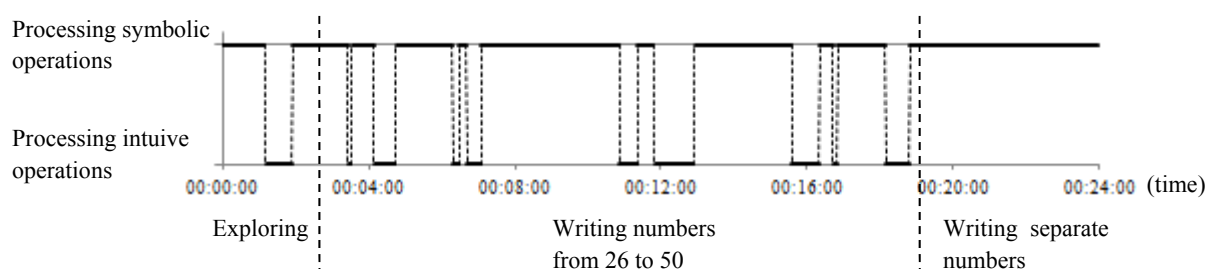


Figure 3. Symbolic and intuitive operations in Mathilde's course-of-action

Mathilde alternated symbolic and intuitive process during the entire first (exploring) and second (writing numbers 26 to 50) phases. For example, after she read the text of the problem, she tried to find a solution in an intuitive way (1:40). She then began to interpret Maya signs with symbolic relations (1:54). Step by step, she constructed right relations between Maya signs and numbers, and started to write Maya numbers after 25. But progressively she leaved this symbolic processing, and

<sup>2</sup> Indications between square brackets are descriptions of the observed behaviours written by the researcher.

<sup>3</sup> At this moment Jonathan didn't know the numbers symbolized by the graphic ensembles he drew: 30sec later (6:11) he had to examine them one after the other in order to identify these numbers.

just copied composed signs from the Maya table and added a dot above everyone. So from 4:08 to 4:58, she wrote two composed signs without knowing the numbers they represent.

As Mathilde repeated this alternation between symbolic and intuitive processes six times, she didn't stabilize symbolic relations between Maya signs and numbers, but changed her interpretation of some elementary signs when she shifted back from intuitive to symbolic processing: for her, a line represented sometimes five and sometimes ten, and a dot up, sometimes ten and sometimes twenty.

## **Discussion**

In these case studies, the same external representations (Maya numerical signs) are alternatively treated in a symbolic way (relying on conventional relations between signs and numbers) and in an intuitive way (relying on visual perceptions of increasing quantities or similarities). It is not surprising that learners use intuitive perceptual cues: as Elby (2000) showed previously for representations in physics, direct and quick processing are allowed by visual attributes, and they may be maintained beyond the acquisition of the ability for a more symbolic processing. Yet our study shows other process than those relying on Elby's "what-you-see-is-what-you-get" (WYSIWYG) general knowledge. During symbolic processing, both subjects identify visual features of the composed signs (increasing quantities and similarities) allowing them further to forget, for a time, the symbolic relations between the sign units (dots and lines) and the numbers. So we can say that both subjects de-symbolize representations. De-symbolization may be seen as a failure: as showed by Mathilde course-of-action, intuitive process may generate incoherence, and hinder the correct achievement of the task. But it may also be seen as functional: as showed by Jonathan course-of-action, intuitive process may allow a quick and easy drawing of a lot of correct signs, and foster the quick achievement of the task. If, as Elby said, intuitive processes are compelling, we can ask if this is because of their functional role as much as compelling visual attributes of the representations.

The de-symbolization in learning Maya numbers could be considered as exceptional because these symbols are not purely arbitrary, but partially based on quantities of signs. Researches with other symbolic representations taught in schools are necessary in order to explore the extension of this phenomenon.

## **References**

- Elby, A. (2000). What students' learning of representations tell us about constructivism. *Journal of Mathematical Behavior*, 19(4), 481-502.
- Theureau, J. (2003). Course of action analysis. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 55-81). Mahwah, NJ: Lawrence Erlbaum.

# How a Picture Can Scaffold Comprehension of Text

Alexander Eitel<sup>1</sup>, Katharina Scheiter<sup>1</sup>, Anne Schüler<sup>1</sup>, Marcus Nyström<sup>2</sup>, Kenneth Holmqvist<sup>2</sup>

<sup>1</sup>Knowledge Media Research Center

Schleichstr. 6, 72076 Tübingen, Germany, a.eitel@iwm-kmrc.de

<sup>2</sup>Lund University, Humanities Lab

Helgonabacken 12, 22362 Lund, Sweden

**Abstract.** We tested the assumption that global spatial information is rapidly extracted from causal system pictures and can be used as a mental scaffold to support mental model construction and thus comprehension of text (*scaffolding assumption*). Students ( $N=84$ ) learned about the structure and function of pulley systems from text or from text with previous presentation of a picture for 600ms, 2sec, or self-paced. Students' eye movements on blank screen while listening to text and comprehension of the pulley system's functioning were assessed. Results revealed that in conditions with initial picture inspection (for 600ms, 2sec, and self-paced) more eye movements in line with the picture's global spatial orientation were made and comprehension was better than in the text-only condition. Results thus suggest that mental model construction from text was (positively) influenced by global spatial information extracted from brief initial picture inspection, thus supporting the scaffolding assumption.

**Keywords:** Learning with text and pictures, eye movements, mental imagery, scaffolding

## Theory

The process of mental model construction is assumed to underlie comprehension of text and pictures (Schnotz & Bannert, 2003). In studies on learning with text and pictures, learners are often required to construct mental models of causal systems, where extracting the system's spatial structure precedes understanding its functions (Hegarty & Just, 1993). Whereas text is in general better suited to express abstract information (e.g., about functions), pictures are better suited to convey spatial information (e.g., spatial structure of the system; Larkin & Simon, 1987). To construct a mental model of the spatial structure of a causal system from text only, text has to be interpreted. This may yield erroneous inferences about the system's spatial structure possibly leading to a mental model that inadequately reflects the contents of the text and thus is detrimental to comprehension of the system's function (Schnotz & Bannert, 2003). In contrast, a causal system picture is analogous to the required structure of a mental model of the causal system, and thus adding a picture to text can facilitate mental model construction (Hegarty & Just, 1993).

Whereas text is usually processed in a sequential-incremental fashion, perception of pictures proceeds from global to local (Navon, 1979). That is, information on a global level (i.e., gist) is extracted from a first glance at the picture, but information on a local level (i.e., visual details) is not. Since the gist of a picture is assumed to represent the picture's global spatial information (Oliva & Torralba, 2006), we assume that the global spatial information extracted from briefly inspecting a causal system picture may act as a mental scaffold, constraining interpretation of subsequent text about the system's spatial structure (Scaife & Rogers, 1996). As a consequence, the process of constructing a mental model of the causal system's spatial structure from text, and in turn comprehension, should be supported by a brief initial glance at the picture (i.e., *scaffolding assumption*). This was investigated in the following using a pulley system as to-be-learned content.

## Prior Study

We conducted a prior study to ascertain that spatial information on a global but not on a local level was extracted from a pulley system picture (see Figure 1). In this study, 85 students (62 female) had to draw a pulley system either without seeing a picture of it, after inspection of the pulley-system picture for 600ms or 2sec, or after self-paced picture inspection time. Results revealed that spatial information on a global level (i.e., orientation of the three pulleys from bottom-left to top-right; see Figure 1) was extracted to a high degree both when the picture was inspected for self-paced ( $M = 28.6\text{sec}$ ) and when

the picture was presented for 600ms and 2sec only. Spatial information on a local level, in contrast, was only extracted when picture presentation was self-paced. Whether the global spatial information extracted from brief inspection of the pulley system (i.e., diagonal pulley orientation) is used to construct a mental model from subsequent text, was investigated in the present study.

## Design and Materials

Students ( $N=84$ ; 60 female) learned about pulley systems in one of four conditions (see Figure 1): (a) text only, (b) a picture presented for 600ms before text (600 before), (c) a picture for 2 seconds before text (2sec before), or (d) a picture for as long as they liked before text (self-paced before). The text first described the spatial structure of the pulley system, followed by an explanation of the function of pulley systems (i.e., what happens when the rope is pulled; cf. Hegarty & Just, 1993). The text contained information about the pulley system's components and their relations; however, it *did not contain any information about the diagonal orientation* of the three pulleys of the system, namely that the upper pulley is located to the upper-right of the middle pulley, and the middle pulley to the upper-right of the lower pulley. Rather, the text described the upper pulley merely to be located *above* the middle pulley, and the middle pulley to be located *above* the lower pulley. Information about the diagonal orientation of the three pulleys thus could be extracted only from the picture. Text was presented auditory while students looked at a blank screen.

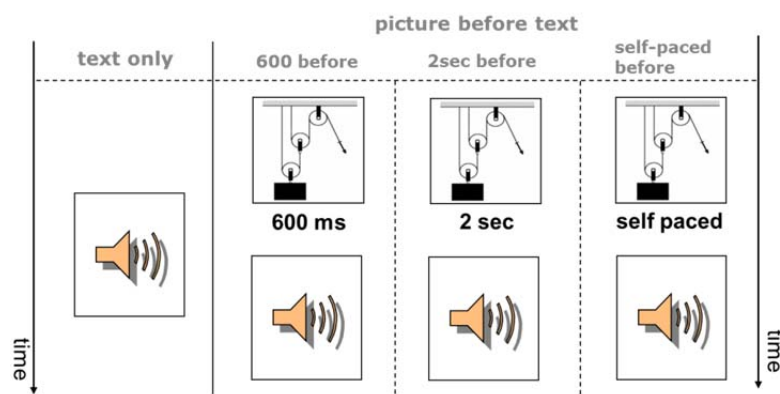


Figure 1. Study Design

## Measures and Procedure

Students' eye movements were recorded while they were listening to text and looking at a blank screen. According to literature on mental imagery, eye movements on blank screen reflect the spatial properties both of a concurrently described scene and of a previously seen visual scene (Johansson, Holsanova, & Holmqvist, 2006). Since in our study the text did not mention a diagonal pulley orientation, more eye movements in the diagonal direction of pulleys as depicted in the picture (i.e., bottom-left to top-right) would indicate that mentally imagining spatial information from text (i.e., constructing a mental model) was influenced by global spatial information extracted from the previously inspected picture. To be able to compare eye movements between experimental conditions, we computed the standardized relative frequency of saccades made in the correct (i.e., according to pulley orientation in picture) versus incorrect (i.e., opposite) diagonal direction.

After learning from text (and picture), students' comprehension of the pulley system's functioning was assessed with a verbal multiple-choice test and a labelling test. Results from both tests were merged into a single score. Multiple-choice items required students both to mentally animate the system's spatial structure (e.g., "if the free end of the upper rope is let go, then the middle pulley turns clockwise"; *yes* was correct), and to reason based on the principle underlying pulley systems (e.g., "if

the weight was attached at the middle pulley, then the rope would have to be pulled with the same force than when the weight is attached to the lower pulley”; *no* was correct).

## Results

Pictures were studied for  $M = 21.5$  sec in the self-paced condition. Eye movements were analysed by including the direction of the standardized relative frequency of saccades (correct vs. incorrect) as a repeated-measures factor in a 2x4 mixed ANOVA with experimental condition (text only vs. 600 before vs. 2sec before vs. self-paced before) as between-subjects factor. Results revealed a main effect of direction,  $F(1, 80) = 23.46$ ,  $p < .001$ ,  $\eta_p^2 = .23$ , and no main effect of experimental condition,  $F < 1$ . Importantly, results revealed a significant interaction,  $F(3, 80) = 2.94$ ,  $p = .04$ ,  $\eta_p^2 = .10$ . Post hoc tests revealed that in the text-only condition, an equal number of saccades were performed in the correct as in the incorrect direction on the blank screen while listening to the text. In all three conditions with picture before text (for 600ms, 2sec, or self-paced), in contrast, more saccades were made in the correct than in the incorrect direction (all  $p_s < .05$ ), suggesting that mental model construction from text was influenced by global spatial information extracted from previous picture inspection.

Comprehension was analysed by means of a one-way ANOVA with experimental condition (text only vs. 600 before vs. 2sec before vs. self-paced before) as between-subjects factor. The ANOVA revealed a significant main effect of condition,  $F(3, 80) = 3.59$ ,  $p = .02$ ,  $\eta_p^2 = .12$ . As expected, post hoc tests revealed that comprehension was better in all three conditions with picture before text (for 600ms, 2sec, or self-paced) compared to the text-only condition (all  $p_s < .05$ ).

## Conclusions and Implications

Global spatial information from the picture was extracted from brief initial inspection (i.e., for 600ms and 2sec). It was integrated with text, supporting the process of mental model construction, and in turn, comprehension. Results are thus in line with the scaffolding assumption.

The present results therefore suggest that to better understand on a fine-grained level how pictures are (best) used in combination with text in a multimedia instruction, it may make sense to first have a closer look at how pictures are processed in isolation and which functions different levels of processing may play during learning.

## References

- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717-742.
- Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science: A Multidisciplinary Journal*, 30, 1053-1079.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Oliva, A., & Torralba, A. (2006). Building the gist of the scene: The role of global image features in recognition. *Progress in Brain Research*, 122, 23-36.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141-156.

# Graphs in Science Education – German Biology Teachers’ Beliefs, Instructional Practices, and Related Student Outcomes

*Carolin Enzingmüller*

University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany,  
carolin.enzingmueller@uni-potsdam.de,

*Claudia Nerdel*

TUM School of Education, Schellingstraße 33, 80799 München, Germany,  
claudia.nerdel@tum.de

*Helmut Precht*

University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany,  
helmut.precht@uni-potsdam.de

**Abstract.** Interpretation and construction of graphs are integral elements of learning science. Research indicates that acquiring these skills is quite a difficult task for learners. Thus, there is a demand for instructional support for students in developing the abilities to interpret and construct graphs. Research indicates that beliefs may be one determining factor in teachers’ instructional practices. This study investigates the relations between teachers’ pedagogical beliefs, instructional practices and student outcomes in the domain of graphing. To measure teachers’ beliefs about teaching and learning graphing, we constructed a questionnaire specifically for the present study. In a pilot study the scales of the teacher questionnaire showed satisfying discriminant validity and reliability. Teachers’ instructional behaviours and student outcomes will be assessed by means of student ratings.

**Keywords:** Teachers’ beliefs, graph-related instruction, biology education

## Theoretical Background and Purpose

Scientists rely on a variety of visualizations to construct and communicate their knowledge (Lemke, 1998). These representations are also present in the science classroom with great frequency (Gilbert, 2007; Roth, Bowen & McGinn, 1999). The interpretation of graphs, diagrams and so forth is therefore clearly an integral element of learning science. But as emphasized by Ainsworth (2011), becoming proficient in science also requires learners to construct their own visualizations. However, acquiring these skills is quite a difficult task for learners (Kozma & Russell, 2005; Schnotz & Bannert, 2003). Studies show that many students have difficulties in interpreting as well as constructing graphs and hold various misconceptions in the area of graphing that may hinder their comprehension (Mautone & Mayer, 2007; Mevarech & Kramarsky, 1997). Thus, there is a demand for instructional support for students in developing the abilities to interpret and construct graphs.

Factors that influence teachers’ decisions regarding instruction are complex and numerous. Research indicates that beliefs may be one determining factor in teachers’ practices (e.g. Mansour, 2009; Pajares, 1992). In the domain of text-picture-integration, Schroeder et al. (2011) showed that teachers’ beliefs affect students’ engagement to learn from texts with instructional pictures. Teachers’ beliefs that students should be taught clear strategies on how to learn from texts with instructional pictures were positively related to students’ self-reported engagement. In contrast, teachers’ beliefs that students should learn to interpret these materials independently were negatively related to students’ self-reported engagement. A mediator analysis revealed that the effects of teachers’ beliefs on students’ engagement were mediated by teachers’ instructional behaviours.

The study described in this synopsis examines German biology teachers’ beliefs and instructional practices involving graphs and their influence on student outcomes. It is guided by the idea that educational beliefs affect decisions about classroom practices, including the area of graphing. The following research questions were addressed:

- 1) What beliefs do German teachers hold about reading and constructing graphs in biology class?



## 2) How are these beliefs related to classroom practices and related student outcomes?

### Methods

A questionnaire was developed in order to gather information from teachers about their educational beliefs in the domain of graphing. Items were modified from an existing questionnaire assessing teachers' pedagogical beliefs in the domain of text-picture comprehension (Schroeder et al., 2011). We also added additional items based on teachers' responses in a previous interview study which targeted general views on teaching and learning graphics in biology class. The following scales were employed: *Importance of practice* was concerned with the belief that the ability to read/construct graphs is important for learning biology. *Independent learning* referred to the belief that students should develop their own interpretations of graphs. *Strategy use* measured the belief that students should learn clear strategies for reading/constructing graphs. *Critical views* was concerned with the belief that students should acquire a critical attitude toward graphs. *Emotional distance* measured if teachers feel insecure when dealing with graph reading/constructing in biology class. *Self-efficacy* assessed the degree to which teachers express optimistic self-beliefs to support students in reading/constructing graphs in biology class.

Our questionnaire comprised 40 items to be rated on a 4-point Likert-type scale (ranging from 1=*strongly disagree* to 4=*strongly agree*). In a pilot study the questionnaire was administered in four teacher training courses. In total, 46 teachers (48% female, 52% male) aged 25 to 64 ( $M=41$ ,  $SD=12.2$ ) participated. All teachers taught biology and chemistry in different urban and rural secondary schools (academic track, *Gymnasium*) in Bavaria.

### Findings

Table 1 provides reliabilities and intercorrelations of the teacher belief scales based on the pilot sample.

Table 1: Reliability estimates (Cronbach's  $\alpha$ ) and intercorrelations of the teacher belief scales ( $N=46$ ).

Scale (number of items)	$\alpha$	Intercorrelations				
<i>Reading graphs:</i>		1	2	3	4	5
1. Importance of practice (4)	.68					
2. Independent learning (4)	.65	.01				
3. Strategy use (4)	.85	.53**	.14			
4. Emotional distance (4)	.75	-.16	-.03	-.16		
5. Self-efficacy (4)	.71	.11	-.10	.22	-.47**	
6. Critical views (4)	.76	.29	.25	.34*	.10	.17
<i>Constructing graphs:</i>						
1. Importance of practice (4)	.84					
2. Strategy use (4)	.87	.49**				
3. Emotional distance (4)	.84	-.39*	-.29			
4. Self-efficacy (4)	.76	.36*	.48**	-.52**		

The Cronbach's  $\alpha$  coefficients are satisfactory and range from .65 to .87. The results of the correlation analysis suggest that there is a reasonable positive association between *importance of practice* and *strategy use*, suggesting that when teachers emphasize the importance of giving students explicit practice in reading and constructing graphs, they are also likely to emphasize that students should learn clear strategies for both, reading and constructing graphs. The correlation analysis also

reveals the theoretically expected association between *self-efficacy* and *emotional distance*. The degree to which teachers considered themselves competent in teaching students to read and construct graphs is negatively correlated with the feeling of insecurity when dealing with graph reading and constructing in biology class. However, most of the correlations between the scales are low to moderate, indicating that they measure independent constructs.

### **Future Steps**

Teachers' instructional behaviours and student outcomes will be assessed by means of student ratings. We are now focusing on designing and testing the student questionnaire.

We would like to present and discuss first findings of the study at EARLI SIG 2 2012.

### **Bibliography**

- Ainsworth, S. E., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333, 1096-1097.
- Gilbert, J. K. (2007). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization in Science Education*. Dordrecht: Springer.
- Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualizations in Science Education* (pp. 121-146). Dordrecht: Springer.
- Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading Science*. London: Routledge.
- Mansour, N. (2009). Science teachers' beliefs and practices: Issues, implications and research agenda. *International Journal of Environmental & Science Education*, 4(1), 25-48.
- Mautone, P. D., & Mayer, R. E. (2007). Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, 99, 640-655.
- Mevarech, Z., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, 32(3), 229-263.
- Pajares, M. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 3, 307-332.
- Roth, W.-M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching*, 36, 977-1019.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141-156.
- Schroeder, S., Richter, T., McElvany, N., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H., & Ullrich, M. (2011). Teachers' beliefs, instructional behaviors, and students' engagement in learning from texts with instructional pictures. *Learning and Instruction*, 21, 403-415.

# Signals in Expository Prose: Is the Ability to Recognize and Interpret them Specific of Reading Comprehension?

*Carlos Ferreira*

"Aravalle" Secondary and Technical School, Spain  
acarlosferreirad@gmail.com

*Hector Garcia Rodicio*

Universidad de Cantabria, Spain.  
Facultad de Educacion - Avenida de los Castros s/n 39005 Santander, Cantabria - Spain  
hector.garciarodicio@unican.es

*Emilio Sanchez*

Universidad de Salamanca, Spain  
esanchez@usal.es

**Abstract.** Expository texts include signals, which guide processing if readers are able to recognize and interpret them correctly. Here we hypothesized that it is easier for readers to process signals presented in spoken modality than to do so with written signals. This is because spoken signals are more expressive and eloquent, which can help readers to recognize and interpret the signals. In an experiment 37 middle-school students read texts with either written or spoken signals and generated a summary. They benefited more from signals presented orally than from signals presented in written modality, as summaries were more complete and organized in the spoken signals condition. To the extent written signals require readers especial abilities to recognize and interpret them, this means that reading comprehension may be more than decoding plus listening comprehension.

**Keywords:** signaling devices, reading comprehension, structural knowledge, Simple View of Reading

## Signals and structural knowledge

Expository texts usually include both information and signals, which are writing devices used by authors to emphasize important information (e.g., “Let me stress that”) and text organization (e.g., “Secondly”; Lorch, 1989). Signals can guide readers' processing, thus enhancing retention and comprehension (Lorch, 1989). However, readers have to recognize and interpret signals if they are to benefit from them. The ability to do so has been called structural knowledge (Goldman & Rakestraw, 2000). The more structure knowledge readers have, the more they take advantage of signals, thus comprehending the text better (Sánchez & García, 2008). In the study reported here we hypothesized that signals presented in spoken modality may be easier to recognize and interpret for readers with low structural knowledge, relative to written signals. This is due to the specific features of spoken signals.

## Pilot study

In order to identify specific features of written and spoken signals we conducted a pilot study. We collected written and spoken discourse that used the same genre (expository), covered the same topics (e.g., Mediterranean Sea), and addressed the same audience (middle-school students). Then, we coded content as either information or signals (i.e., all devices that could be deleted from discourse without affecting information; Lorch, 1989). We focused on a specific kind of signals: enumeration devices.

We found two main differences between written and spoken devices. They differ on *expressiveness*. Spoken ones are expressive, since they are marked by prosody (e.g., “The S-E-C-O-N-D R-E-A-S-O-N is”) whereas written ones are flat (e.g., “The second reason is”). Moreover, devices differ on their *eloquence*. Spoken ones are eloquent, i.e., they provide hints to readers about how to interpret them (e.g., “Now we will consider a second idea, which is different from that explained before but which is also very important”); conversely, written ones are concise (e.g., “Secondly”).

## Experiment

We hypothesized that expressiveness and eloquence of the spoken modality would help readers with low structural knowledge to recognize and interpret the signals in a text. Low-structural-knowledge readers receiving spoken signals should comprehend the text better than those receiving written ones.

### *Participants and design*

Eighteen middle-school students (10-12 years old) served in the *written signals condition* and 19 served in the *spoken signals condition*. We chose participants of this particular grade because we expected them to have low structural knowledge (this was confirmed; see below).

### *Procedure and materials*

First participants solved a *structural knowledge test* (Sánchez & García, 2008). It consisted of 12 incomplete microtexts about an imaginary country (Paíta), for which participants had to find a completion by correctly interpreting the signal in the text (scores 0-12). Cronbach's alpha was .81. Example: "Paíta is a very rich country. This is due, firstly, to the amount of oil and other minerals it has." A good completion was one including "secondly", "moreover", or a similar enumeration device.

Afterwards, participants took one task of a standardized *reading comprehension test*. They read two short, descriptive texts (Papuan people, the Inuit) and solved 10 open-ended questions afterwards (scores 0-10). These texts do not include signals, thus they require no structural knowledge to readers.

Then, participants read a *text* on the Mediterranean Sea, explaining the three causes of its death. It consisted of 300 words and was presented in written modality on a computer screen. The text also included enumeration signals that proved effective in promoting comprehension in a prior study (Sánchez & García, 2008). In the *spoken signals condition* these signals were presented orally and were audio-recorded by a trained speaker. They were designed to be expressive and eloquent (e.g., "After considering ONE REASON, we are going to talk about the S-E-C-O-N-D one, which is also very important. This reason is... "). In the *written signals condition* the signals were presented as on-screen text. These signals were flat and concise ("The second reason is... ").

Therefore, participants in the spoken condition read the information parts of the text and listened to audio recordings containing the signals whereas participants in the written condition read both the information parts and signals. In order to avoid a possible "surprise effect" in the spoken condition, we used two warming-up texts. They both were short (80 words), descriptive texts (Ratel, the Vikings) involving six sentences. In each text, half the sentences were presented in spoken modality and the rest in written modality. After reading the text, participants solved one question (one was about contents in a spoken sentence, another one was about those in a written sentence). By reading these texts and answering the questions, we expected participants to get familiar with the mixed spoken-written format and, other things being equal, to perceive spoken and written contents as equally important. Accordingly, possible differences in the results could not be attributed to a mere "surprise effect".

Finally, participants wrote a *summary* of the text. Quality of summaries depended on the number of central ideas they included and whether they were structured (scores 0-4). Interrater agreement was .90

### *Results*

As shown in Table 1, participants had low levels of structural knowledge and normal reading comprehension (90<sup>th</sup> percentile). As indicated by two independent samples *t*-tests, there were no differences between conditions in these variables ( $p$ 's > .50). An ANCOVA with condition as the between-subjects factor, summary as the dependent variable, and reading comprehension as the covariate revealed a significant effect of condition, with participants in the spoken condition performing better than those in the written condition,  $F(1, 34) = 5.47$ ,  $MSE = 0.88$ ,  $p = .03$ .

Table 1: Performances of all conditions in all variables.

	Written signals condition		Spoken signals condition	
	M	SD	M	SD
Structural knowledge	5.78	2.32	6.21	3.69
Reading comprehension	5.94	3.19	6.68	3.51
Summary	1.28	0.67	2.11	1.29

## Discussion

To benefit from signals, readers have to recognize and interpret them. Signals in spoken modality are expressive and eloquent, which may help readers to perform the processes of recognition and interpretation. Here we confirmed this hypothesis: readers with low structural knowledge generated more complete and organized summaries from texts that included spoken signals, relative to a written signals condition. Future research could explore the specific role of expressiveness and eloquence.

One might argue that the spoken advantage is due to the fact that participants in this study had not mastered the process of decoding. Given so, participants found it more difficult to process written words than spoken words, regardless of the role of the words (i.e., informing or signalling). However, this possibility does not seem plausible, as participants exhibited reading comprehension scores within normal ranges, as measures by a standardized test (that included texts without signals, which then demand no structural knowledge on the part of readers). Poor decoding would have affected this score.

Furthermore, results do not derive from a “surprise effect” due to the switch from one modality to another in the spoken signals condition. In order to avoid this potential effect, we conducted a warming-up phase in which participants of both condition got used to such mixed format.

If written signals are more difficult to process than spoken signals, it means that reading comprehension of signalled prose requires readers to have specific knowledge beyond decoding, as compared with listening comprehension. This interpretation can be viewed as a revision of the Simple View of Reading, which establishes that reading comprehension equals decoding plus listening comprehension (Gough & Tunmer, 1986).

## References

- Goldman, S. R., & Rakestraw, J. A. (2000). Structural aspects of constructing meaning from text. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (pp. 311–335). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.
- Lorch, R. F. (1989). Text signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, 1, 209-234.
- Sánchez, E., & García, R. (2008). The relation of knowledge of textual integration devices to expository text comprehension. *Reading and Writing*, 22, 1081-1108.

# On-Line Assessment of Students' Global Reading Strategies through Triple Task Technique

*Raquel Fidalgo*

Universidad de León. Departamento de Psicología, Sociología y Filosofía; Campus de Vegazana s/n, 24071  
León, Spain  
rfidr@unileon.es

*Olga Arias*

Universidad de León. Departamento de Psicología, Sociología y Filosofía; Campus de Vegazana s/n, 24071  
León, Spain  
oorig@unileon.es

*Mark Torrance*

Nottingham Trent University. Division of Psychology; Nottingham, United Kingdom  
mark.torrance@ntu.ac.uk

*Thierry Olive*

Centre National de la Recherche Scientifique; Poitiers, France.  
thierry.olive@univ-poitiers.fr

*Rui A. Alves*

Universidade do Porto. Faculdade de Psicologia e de Ciencias da Educaçao; Porto, Portugal  
ralves@fpce.up.pt

**Abstract.** We present an on-line assessment method named the Triple Task which allows us to estimate the global reading strategies used by students in a normal reading task at schools when reading and summarizing a text. This technique provides both an estimation of the cognitive costs of different actions and processes through the reading task, and an estimation of the relationship between readers' global strategies and the performance scores of reading comprehension. First, we describe the general design of the Triple Task and its adaptation for the reading research field. We discuss its application and the possibilities it offers for doing research in the reading field as well as the potential limitations of the technique. Second, we assessed the possible reactivity of the Triple Task on reading comprehension in different age groups. Findings supported the conclusion that the triple task is a non-reactive method of assessment in reading.

**Keywords:** On-line assessment of reading process, Reading strategies, Cognitive effort, Reading comprehension, Summary

## Design of the Triple-Task Technique on Reading Research

The triple task technique arose from the cognitive research in the writing research field as a variation of the Secondary Task technique. Although it has been widely used to study a considerable number of issues in relation to the writing research field (Olive, 2009; Olive, Kellogg, & Piolat, 2002), it has not been used frequently to investigate text reading strategies. Generally, it has only been used in studies focused on reading but in the context of sub-processes of revision process of writing.

As its name indicates, it includes three different tasks that the subjects have to perform. The primary task is the focus of research interest. We decided to use as primary task, a very common and usual reading task for learning in educational contexts, such as, reading to summarize and comprehend an expository text. Students had to read and summarize an expository text in a paper and pencil task. We used two expository texts that belong to a standardized comprehension test in Spain (Vidal-Abarca, Gilabert, Martínez, & Sellés, 2007). Such expository texts were appropriate for students' ages, knowledge and interests and they were counterbalanced across the grades and students' genre.

The second and third tasks were administered by specific computer software that was designed specifically for the purposes of this reading comprehension research. Students in the second task had to detect an auditory signal and respond to it by saying aloud as soon as possible "ta". To do so, they used individual headphones and microphones. The assumption of this second task is that the time span taken to detect the signal and give the specific response is an indicator of the cognitive load that

the student experienced during the activity or process that they were implementing when the signal sounded. Before this task, students were trained and we measured the baseline of students' reaction times.

The third task consisted of a direct retrospection about what students were doing when they detected the auditory signal, using a set of specific categories of retrospection. The design of the different categories of retrospection in the reading task took into account a suitable number of categories that were easily differentiated by younger students so as to ensure reliable categorizations by students. In addition, these categories allowed us both to record interesting data about the global reading strategy used by students to move through the text and summarize it, and to estimate the cognitive cost or load of processing involved in reading different pieces of text. Taking into account these assumptions we designed six categories of retrospection: reading the text, thinking about the text or the summary, taking notes in the text, writing a summary, reading the summary and others. Besides, the reading text category includes different retrospections according to the piece of text that students were reading in each moment.

In general, the design of the triple task technique on this reading research context will provide interesting data for doing research on reading. The analysis of the global distribution or sequence of activities through the reading task provides a description of the global comprehension strategies of students in a common reading task in schools. This allows us to investigate how different samples of students (i.e. skilled and less skilled, students with and without reading disabilities or students of different ages) differ in (i) their global reading strategies; (ii) self-regulation of actions; and (iii) interactions with the text and task through the reading process. Besides, students' performance in summaries task and in the reading comprehension test will make possible to examine if some students' global reading strategies are associated with different performance scores. At the same time, changes in reaction time measures indicate the level of students' involvement in the activities and processes of the main task. Furthermore, the degree of effort or cognitive load associated with the reading activity of each segment of text suggests the cognitive cost of readers' processing concerning each particular segment of text. On the other hand, some potential problems of the triple task technique have been suggested in relation to its reactivity. The main criticism about the triple task technique is if it disrupts the main task of interest. Several previous researches in the writing field have concluded that the triple task technique is not reactive (see Olive et al., 2002), however, from our own knowledge there are not empirical data about the possible reactivity of the triple task in the field of reading research; purpose of the present study.

## **Method**

### *Participants*

439 Spanish students distributed in 5<sup>th</sup> (87 students) and 6<sup>th</sup> (96 students) grades of Primary Education and 1<sup>st</sup> (124 students) and 2<sup>nd</sup> (132 students) grade of Secondary Education formed the triple task group. 109 Spanish students of the same grades (32 students of 5<sup>th</sup> grade; 32 students of 6<sup>th</sup> grade, 21 students of 1<sup>st</sup> grade and 24 students of 2<sup>nd</sup> grade) were the control group, that is, they implemented the reading and summarizing task in isolation without the triple task technique.

### *Measures*

*Reading Comprehension Achievement.* We measured the performance of both groups in reading comprehension previously through a Spanish standardized test of reading comprehension that includes literal and inferential comprehension questions and cloze test. These measures of reading achievement were used as covariates.

*Reading Comprehension Test and Summary.* After the triple task assessment, students had to implement a reading comprehension test on computer in relation to the summarized text to get a measurement about students' performance in reading comprehension. This test included 10 multiple-choice questions of each text that were designed based on Kintsch's comprehension model (Vidal-Abarca et al., 2007). The questions were divided into four types: text-based, bridging inferences, global inferences and elaborative inferences.

As for the summaries, we scored students' text summaries to get an estimation of how well they had constructed a right text-based representation, scoring the number of type of ideas of text included in summaries according to a previous analysis of the texts by different experts.

## Results and Conclusions

We made MANOVA statistical analyses taking triple task condition as independent variable, previous reading achievement as covariates, and summary scores and a global score of reading comprehension text as dependent variables. Results showed there were not statistically significant differences in comprehension and summaries scores between students of the triple task and control groups ( $F(5, 541) < 1$ ), and also there was no evidence of an effect of the interaction between covariates and triple task condition.

Table 1: Descriptive Measures of Summary and Text Comprehension Scores

	Triple Task Condition	Control Condition
Summary scores		
Number of ideas of level 1	14.53 (5.58)	13.65 (5.67)
Number of ideas of level 2	7.19 (3.47)	6.64 (3.44)
Number of ideas of level 3	4.94 (2.70)	4.28 (2.38)
Number of ideas of level 4	1.28 (1.29)	1.00 (1.18)
Total score of reading comprehension test	6.12 (2.23)	6.09 (2.22)

Findings supported that the triple task is a non-reactive method of assessment in reading, at least in these groups of age, which will allow us to study the reading comprehension process of students from different points of view.

Note. This research was possible with funds from Science and Innovation Ministry of Spain through research project (Reference: EDU2010-18219), that was awarded to Dra. Fidalgo.

## References

- Olive, T. (2009). Methods, techniques, and tools for the on-line study of writing process. In N. L. Mertens (Ed.), *Writing: processes, tools and techniques (1-18)*. New York: Nova Science Publishers.
- Olive, T., Kellogg, R. T., & Piolat, A. (2002). The triple task technique for studying the process of writing. En T. Olive, & C. M. Levy (Eds.) *Contemporary Tools and Techniques for Studying Writing* (pp. 31-60). Dordrecht, The Netherlands: Kluwer Academic Press.
- Vidal-Abarca, E., Gilabert, R., Martínez, T., & Sellés, M. P. (2007). *Test de estrategias de comprensión*. Madrid: Instituto Calasanz de Ciencias de la Educación.



# The Effect of Task Instruction and Text Overlap on the Integration of Multiple Cross-Domain Sources

Carla Firetto, Peggy Van Meter  
228 Cedar Building, The Pennsylvania State University  
University Park, PA, USA  
Cmf270@psu.edu, Pnv1@psu.edu

**Abstract.** In this study, we examined the impact of task instruction and surface feature overlap between texts. All participants read two related cross-domain texts, one from psychology the other from statistics. Texts were manipulated to have either explicit overlap or implied overlap of surface features. Participants were also given task instructions prior to reading the texts that either prompted them to read to integrate or read to comprehend. Outcome measures included a task verification measure of participants' recall of the instructions, a single-text comprehension test, and an integration inference verification test. Neither of the condition manipulations affected the comprehension of individual texts. Despite participants' ability to correctly verify knowledge of the assigned task, instruction type delivered did not affect integration. Participants reading the pair of texts with explicit overlap scored higher on the integration measure than those reading the texts with only implied overlap.

**Keywords:** multiple sources, intertextuality, text integration, task instructions

## Introduction

Learners of all ages are increasingly presented with multiple text sources. An often unspoken expectation is that readers will integrate and connect between these texts. While research on learners' integration of multiple sources is still developing, existing research suggests that for many readers, integration between texts is not spontaneous (Wineburg, 1991; Wolfe & Goldman, 2005).

Theoretical and empirical research on single-text reading comprehension underlies the research on multiple texts. Kintsch's (1988) CI Theory provides a strong framework of single-text comprehension and is used as a foundation for multiple-text comprehension. The newest revision to the documents model, MD-TRACE, (Rouet & Britt, 2011) emphasizes the importance of task instructions and information resources (i.e., the texts) on integration and provides the framework for this study.

Past research by the authors (Firetto & Van Meter, 2008) tested the influence of two factors—task instruction and text overlap—on single-text comprehension, problem solving, and integration. A priori task instruction, either to *comprehend* or to *integrate*, had no effect on problem solving or integration. When participants read a pair of texts manipulated to provide *explicit overlap* of surface features, compared to a pair of texts that had only *implied overlap*, greater problem solving and integration was evidenced. Single-text comprehension did not differ regardless of task instructions or text overlap.

The study presented here enhanced task instructions and added a measure to verify participants had knowledge of the task (i.e., read the instructions). Finally, we tested whether the findings would hold when using an alternate measure of integration (i.e., inference verification).

## Method

### *Design and Participants*

All participants read two related texts, one from psychology and one from statistics, with reading order counterbalanced across participants. Participants were randomly assigned to conditions in a 2 (task instruction) x 2 (text overlap) factorial design. Participants were 115 undergraduate students (85 female; 111 between 18-25 years old) at a large university in Northeastern United States.

### *Materials*

**Task Instructions.** Participants were instructed to either *comprehend* or *integrate* the texts. Participants in the comprehend condition were told to “read and understand the texts”. In the integrate condition,

participants were to “think about how each text relates to the other” as well as to “understand how information from both can be combined” and that doing so could lead to greater understanding.

*Text Materials.* Participants read a pair of texts, one explained the concept of correlation (statistics text; 2,300 words; Flesch reading ease, 37) and one described the relationship between mental illness and intelligence (psychology text; 1,600 words; Flesch reading ease, 27) with either *implied overlap* or *explicit overlap*. The implied overlap condition had no explicit connection between the texts; connections were only inferred. The explicit overlap condition had manipulations to the words, examples, and a visual representation to provide overlap between texts. The implied overlap psychology text used the words “link” and “connection” to refer to the association between mental illness and intelligence; the explicit overlap psychology text replaced those words with “correlation” and “relationship” (exact words used in the statistics text). Statistics text examples in the implied overlap text were non-human (e.g., temperature); the examples in the explicit overlap text were human (e.g., reading ability). All representations in the implied overlap psychology text were pictorial; the explicit overlap version replaced one with a scatterplot (representation used in the statistics text).

*Task Instruction Verification.* Participants were asked to restate the task instructions after delivery and before presentation of the texts. Each response, irrespective of condition, was coded as either *comprehension based* (i.e., “read” or “understand”) or *integration based* (i.e., “relate” or “combine”).

*Intratext Posttest and Intertext Posttest.* The *intratext* posttest contained 20 multiple-choice single-text comprehension questions; distributed equally between each text. The *intertext* posttest contained 15 statements that could either be reasonably inferred from integration of the texts or not. Each of these statements contained content from both texts, and the item order was randomized (c.f., Royer, 1990).

### *Procedures*

All materials were completed electronically in a computer lab. After consent, participants were randomly assigned to one of four conditions. All participants (1) completed the demographic questionnaire, (2) read assigned task instructions, either to *integrate* or *comprehend*, (3) responded to the task instruction verification item, (4) read both texts, with either *implied* or *explicit* overlap, (5) responded to the intratext posttest items, and finally, (6) responded to the intertext posttest items.

## **Results**

### *Effect of Task Instruction*

A Chi-Square test of independence tested the relationship between the instructions delivered to the participants and the participants’ ability to recall these instructions later (Table 1). Yates’ correction was calculated to account for the binomial distribution. The association between condition and participants’ responses to the instruction verification task was significant,  $\chi^2(1, N = 113) = 93.77, p < .05$ . To measure strength, the phi coefficient ( $\phi$ ) was used as a measure of effect size,  $\phi = .93$ . As Table 1 shows, there is evidence that participants did read and recall condition instructions.

### *Intratext and Intertext Comprehension*

The reliability for the 20-item intratext comprehension test was .570. A 2 (task instruction) x 2 (text overlap) ANOVA found no main effect for either text overlap,  $F(1, 109) = .004, p > .05, \eta^2 = .000$ , or task instruction,  $F(1, 109) = .294, p > .05, \eta^2 = .003$ ; nor was the interaction significant,  $F(1, 109) = .537, p > .05, \eta^2 = .005$ . These means are located in Table 2.

The reliability for the 15-item intertext comprehension test was .268. Table 2 contains condition means for this variable. We predicted main effects for task instruction to integrate, explicit text

overlap, as well as an interaction between the two forms of integration support. This prediction was tested in a 2x2 ANOVA with task instruction and text overlap as independent variables. The main effect for text overlap was significant,  $F(1, 109) = 4.382, p < .05, \eta^2 = .039$ , but there was no main effect of task instruction,  $F(1, 109) = .453, p > .05, \eta^2 = .004$ . The interaction was nonsignificant,  $F(1, 109) = .383, p > .05, \eta^2 = .004$ .

Table 1: Number of participants' comprehension vs. integration coded response on the task instruction verification item by instruction delivered.

	Comprehension Coded Response	Integration Coded Response
Instruction to Comprehend	61	1
Instruction to Integrate	3	48

Table 2: Means and standard deviations of posttest measures according to condition.

	Instruction to Comprehend		Instruction to Integrate	
	Implied Overlap	Explicit Overlap	Implied Overlap	Explicit Overlap
Intratext	13.93 (2.97)	14.29 (2.62)	14.03 (3.06)	13.60 (2.64)
Intertext	10.35 (1.92)	9.79 (2.00)	9.80 (2.21)	10.85 (1.79)

## Conclusions and Limitations

Results from the intratext comprehension measure suggested that the neither of the condition manipulations, intended to influence integration of the texts, affected comprehension of individual texts. However, individuals reading a pair of texts with explicit overlap scored higher on a measure of intertext comprehension than did those reading the texts that lacked explicit overlap. Additionally, the type of instruction delivered did not affect performance on this intertext comprehension measure, even though participants were able to successfully verify that they had knowledge of their assigned task. These results replicate findings from our previous study (Firetto and Van Meter, 2008).

A major limitation of this study is the low reliability of the intertext comprehension measure. As a result of this low reliability, the impact of the task instructions condition may have been masked; however, the potency of the text overlap condition still results in significant differences despite increased error from low reliability. Practical significance is also evident as these results replicate findings from our previous study (Firetto & Van Meter, 2008) using a different measure of integration.

## References

- Firetto, C. M., & Van Meter, P. (2008). The effects of surface features and reader's goals on intertextuality. Paper presented at the Annual Meeting of the American Educational Research Association. NY, NY.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*, 95(2), 163-182.
- Rouet, J. F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M.T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text Relevance and Learning from Text* (pp.19-52). Charlotte, NC: Information Age Publishing
- Royer, J. M. (1990). The Sentence Verification Technique: A new direction in the assessment of reading comprehension. In S. Legg & J. Algina (Eds.), *Cognitive Assessment of Language and Math Outcomes* (pp. 144-191). Norwood, NY: Ablex Publishing Corporation.
- Wineburg, S. (1991). Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73-87.
- Wolfe, M. B. W., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, 23(4), 467-502.

# Critical Aspects of Scientific Phenomena – to the Fore, in the Dackground, or not Present in Scientific representations

*Tobias Fredlund, John Airey, Cedric Linder*  
Uppsala University  
Box 516, 850 20 Uppsala, Sweden  
tobias.fredlund@physics.uu.se

**Abstract.** In order to talk about how meaning can be construed from scientific representations we draw on the phenomenological term appresentation. Appresentation refers to those parts of an object that are not readily presented, but experienced as co-present with the presented object. We review a number of scientific representations that students have had reported difficulty with using in certain prescribed situations. Our analysis indicates that representations that are conventionally used in many scientific situations often do not lead students to appresenting the critical aspects of scientific phenomena. Critical aspects of scientific phenomena can range from being present and foregrounded (salient), through present but in the background, to not present at all in the representation at hand. Our discussion suggests that pedagogical implications of these conclusions include the need for teachers to be aware of that their taken-for-granted interpretation of common representations may not be the same as students' interpretations, especially in unfamiliar or less common situations.

**Keywords:** Representations, appresentation, awareness, science, physics, critical aspects

## Introduction

Scientific communication is often not as transparent as many instructors may expect. For example, Sheila Tobias (1986) invited non-science professors to attend carefully prepared university level physics lectures given by teachers recently awarded prizes for their teaching. One of the attending professors who was interviewed, afterwards said: 'It seemed to me during these lectures that I lacked any framework of prior knowledge, experience or intuition that could have helped me order the information I was receiving. I had no way of telling what was important and what was not.'

Science students experience similar challenges in their classes. For example, Airey and Linder (2009, p. 37) quote a student who, having been exposed to a presumably well-known representation depicting a transformer together with written text on the whiteboard, in a stimulated recall interview commented, '.... I don't know what this is. I didn't know what he [the teacher] was writing...' And: 'it's, quite often like that in the lectures – that he's drawing something on the whiteboard and he assumes that we know this from before.'

In both examples above, the interviewees appeared to expect that what the instructor said and did would make them aware of something they would recognize from before and be able to refer to. In the former example some kind of knowledge structure, and in the latter some details of the laboratory equipment represented in a drawing. Marton & Booth (1997, p. 99) point out a 'highly critical aspect of awareness', which can be useful in order to talk about the experiences of the students in the examples above, namely 'appresentation' (cf. Husserl, 1931, 1973; and Schutz, 1962).

## *Appresentation.*

Appresentation is that, which lies behind the visually experienced and is simultaneously 'co-present' in the experiencing of a presented object. For example, 'the strictly seen front of a physical thing always and necessarily appresents a rear aspect and prescribes for it a more or less determinate content' (Husserl, 1931, 1973, p. 109); The 'appresented is co-being along with what exists' (Husserl, 1989, p.

352). Schutz (1962) calls this relationship ‘appresentational pairing’, where an appresenting member (e.g. the ‘front’) of a pair represents an appresented member (e.g. the ‘rear aspect’)<sup>4</sup>.

Following Airey and Linder (2009), we draw on this idea of appresentation to explore that which is often taken for granted in the intended meaning of representations when being used in educational situations. We present an analysis of examples of representations that are often used in physics. Some of the analysed representations, and the situations they represent, have previously been shown by research to be problematic for many students.

### *Research questions.*

Our research interest as presented above led us to formulate the following research question to guide our analysis: Does a particular representation present all the critical aspects that a learner needs to be aware of in order to explain the represented physical phenomenon, make predictions of the represented physical situations, or solve related physics problems?

### **Method**

The analysed representations were chosen on the basis of students appearing to have problems with working effectively with them. Some of these representations have previously been analysed by others (see references in the analysis section). In these cases we have interpreted the existing analyses through our theoretical lens in order to describe how the notion of appresentation can provide a useful and fruitful way to discuss the learning challenges that they present to the students. In one case, the analysis was made from our own video data. The analysis was accomplished by applying our knowledge as physics teachers in conjunction with using a multimodal discourse (Kress & van Leeuwen, 2001) lens to look for features of the presented representations that may potentially be paired with critical aspects of the phenomena – or for the problems – at hand. A similar analysis, using Lemke’s (1990) thematic patterns, was recently made, where such potential of representations was called ‘disciplinary affordances’ (Fredlund, Airey, & Linder, 2012, p. 658), these were defined as ‘the inherent potential of that representation to provide access to disciplinary knowledge’.

### **Analysis**

Examples of situations where representations were analysed include:

1) Getting a bulb to shine using a battery, a bulb and only one wire (McDermott & Shaffer, 1992; Redish, 2003; Shaffer & McDermott, 1992). Here the interior construction of a bulb has to be appresented; appresenting the bulb as a unit without its parts is not sufficient for constituting an appropriate understanding.

2) Seeing the refraction of light in terms of a changing the speed of light across two different optical media. The speed of light can potentially be perceived as proportional to the distance between wave fronts in a wave front diagram, yet there is nothing directly representing speed in a ray diagram (Fredlund, et al., 2012).

3) Appreciating that the normal force is not always as big as, and in opposite direction to, the force of gravity requires appresenting the earth, and the forces acting on the earth.

4) Being able to build advanced electric circuits in the laboratory means having to appresent those parts that by convention are not represented in circuit diagrams, but often taken for granted (Stetzer, 2011).

---

<sup>4</sup>In many ways, Kress’s (2010, p. 70) analysis of the basis of representation is similar: a ‘relation of analogy’ (this *is like* that) which in turn leads to ‘a metaphor’ (this *is* that) exists between the signifier and the signified.

## Discussion

At a number of places in our analysis there are opportunities for appresentational pairing between features of a representation and the critical aspects of the physical phenomena at hand. Following Fredlund, Airey and Linder (2012) we suggest that these opportunities constitute the disciplinary affordances of the representation. At other places in our analysis, such a pairing between representational features and critical aspects of the phenomenon was either absent, or hardly noticeable. The critical aspects needed must then be inferred and appresented from the represented situation, thus potentially represented separately, juxtaposed with, or inserted into already existing representations, or represented in another format or semiotic system, which may in turn provide the needed ‘disciplinary affordances.’ Conversely, the appresented aspects need also be taken to be critical.

## Conclusions

Our results suggest that a directly visual or passive interpretation of representations is often not sufficient for making appropriate inferences in the different cases mentioned in our research question. Rather, students often require access to information either not presented or not foregrounded in the representations. This is information that students need to appresent on the basis of previous experience of situations that are similar to the situation at hand. Pedagogical implications of these conclusions include the need for instructors to be aware of the risk that their taken-for-granted interpretation of common representations may not be the same as their students’ interpretations of the same representation—especially in unfamiliar or less common situations. Furthermore, the kind of analysis presented herein may provide instructors with a tool to identify those critical aspects of the studied phenomena that students may lack access to. Such identification may potentially change the ways in which certain representations are used in instruction, and how this use is motivated.

## References

- Airey, J., & Linder, C. (2009). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, 46, 27-49.
- Fredlund, T., Airey, J., & Linder, C. (2012). Exploring the role of physics representations: an illustrative example from students sharing knowledge about refraction. *Eur. J. Phys.*, 33, 657-666.
- Husserl, E. (1931, 1973). *Cartesian Meditations : an Introduction to Phenomenology* (D. Cairns, Trans.). The Hague: Dordrecht : Kluwer.
- Husserl, E. (1989). Collected works. Vol. 3, Ideas pertaining to a pure phenomenology and to a phenomenological philosophy, second book : studies in the phenomenology of constitution (R. Rojcewicz & A. Schuwer, Trans.). The Hague: Nijhoff.
- Kress, G. (2010). *Multimodality : A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal Discourse : The Modes and Media of Contemporary Communication*. London: Arnold.
- Lemke, J. L. (1990). *Talking Science*. Norwood, New Jersey: Ablex Publishing.
- Marton, F., & Booth, S. (1997). *Learning and Awareness*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- McDermott, L. C., & Shaffer, P. S. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *Am. J. Phys.*, 60(11), 994-1003.
- Redish, E. F. (2003). *Teaching Physics : with the Physics Suite*. Hoboken: John Wiley & Sons.
- Schutz, A. (1962). Collected Papers I : The Problem of Social Reality. The Hague: Martinus Nijhoff.
- Shaffer, P. S., & McDermott, L. C. (1992). Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of instructional strategies. *Am. J. Phys.*, 60(11), 1003-1013.
- Stetzer, M. R. (2011). *Upper-division analog electronics: When is a lab course really a lab course?* Paper presented at the Foundations and Frontiers of Physics Education Research 2011, Bar Harbor, Maine.
- Tobias, S. (1986). Peer Perspectives: On the Teaching of Science. *Change*, 18(2), 36-41.

# A Comparison of Prompts, Corrective Feedback, and Elaborative Feedback in Multimedia Learning

*Hector Garcia Rodicio*

Universidad de Cantabria

Facultad de Educacion - Avenida de los Castros s/n 39005 Santander, Cantabria, Spain

hector.garciarodicio@unican.es

*Emilio Sanchez*

Universidad de Salamanca, Spain

esanchez@usal.es

**Abstract.** In one experiment we compared different forms of support for self-regulated learning. Ninety-five undergraduate students with low prior domain knowledge learned plate tectonics from a multimedia presentation involving animation with concurrent narration and support in one of four forms: elaborative feedback on answers to adjunct questions, corrective feedback on answers to adjunct questions, prompts to monitor and revise understanding, and no support. After the presentation, participants solved retention and transfer tests. Results revealed that participants receiving elaborative feedback outperformed those in the other conditions, which did not differ from one another.

**Keywords:** Self-regulated learning, instructional support, corrective feedback, elaborative feedback, prompts, multimedia presentations

## Self-regulation in learning complex conceptual systems

Multimedia and hypermedia materials involve explanations of complex cause-and-effect systems, such as plate tectonics. To acquire this complex conceptual knowledge, learners have to select, organize, and integrate multiple pieces of information; use the relevant prior domain knowledge they have to perform the integration; and do all this within the limited capacity of working memory (Mayer, 2009).

In order to deal with these conditions, learners have to self-regulate their learning. Self-regulated learning is an active process in which learners set learning goals, monitor their progress toward the goals, and try to control themselves in order to make better progress toward the goals (Pintrich, 2004). Carrying out these self-regulatory processes is critical for deep learning from multimedia and hypermedia materials (Azevedo, Guthrie, & Seibert, 2004). However, evidence also indicates that learners find it difficult to self-regulate their learning (Azevedo et al., 2004). This warrants the use of some form of support for self-regulated learning.

## Support for self-regulation in prior research

Researchers have developed ways of supporting learners to self-regulate while learning from multimedia and hypermedia materials. One strategy is to provide learners with *prompts*, which are devices that elicit the execution of self-regulatory processes, such as planning (“Decide what you want to learn”) or monitoring understanding (“Check your understanding of what you have learned so far”). Evidence indicates this is an effective strategy (Azevedo, Moos, Greene, Winters, & Cromley, 2008).

Another strategy is to insert questions into the material and provide *feedback* on learners’ answers. Feedback can consist of either telling learners whether they are right/wrong (corrective feedback) or giving them an explanation of why their answers are right/wrong and of what the correct answer is (elaborative feedback). These feedback messages support learners in monitoring and regulating their ongoing understanding, which are important self-regulatory processes. There is some evidence that feedback works effectively (Lee, Lim, & Grabowsky, 2010).

Prior research has explored ways of supporting self-regulatory processes while learning from multimedia and hypermedia materials finding that both prompts and feedback can be effective. However, prior studies have not compared the effects of different kinds of support with each other.

Our goal was to address this question. Accordingly, we compared the effects of prompts, corrective feedback, and elaborative feedback. In an experiment participants were asked to learn about plate tectonics from a multimedia presentation, which included support in one of the four forms.

## Method

### *Participants and design*

Ninety-five undergraduate students (aged 18-21) participated in the experiment. They were randomly assigned to one of four conditions: elaborative feedback condition ( $n = 24$ ), corrective feedback condition ( $n = 25$ ), prompts condition ( $n = 23$ ), or control condition ( $n = 23$ ).

### *Procedure and materials*

First participants took a *prior knowledge test* on Geology, which consisted of eight open-ended questions. Scores ranged from 0 to 24.

Afterwards, participants studied from a *computer-based multimedia presentation* on plate tectonics that included modules with animation and concurrent narration. Participants went over the modules at their own pace. The modules were designed according to the principles suggested by Mayer (2009).

Depending on the condition, presentations also included *support devices*. In the *elaborative feedback condition* (a) multiple-choice questions were inserted into the presentation. They addressed aspects of plate tectonics that are commonly misunderstood (based on pilot studies). After choosing an answer, participants received (b) a corrective message, which pointed out if their answers were right/wrong. Afterwards, they received (c) an explanation of why their answers were right/wrong and of what the correct answer was. It should be noted that the explanation was based on the contents of the presentation and did not add new contents. In the *corrective feedback condition* participants received the same support devices, (a) and (b), except for the last part, i.e., (c). In the *prompts condition* participants were provided with messages that warned about potential misunderstandings (corresponding to the aspects covered by the questions in the other conditions) and asked them to monitor and revise their understanding. The *control condition* involved no support devices.

Finally, participants solved a *retention test* with six open-ended questions (scores 0-18), which required them to recall key aspects of the presentation. Participants also took a *transfer test* with eight open-ended questions (scores 0-24), which required participants to use the knowledge they had acquired to solve novel problems.

### *Results*

All performances are shown in Table 1. A MANCOVA with condition as the between-subject factor and prior knowledge as a covariate on retention and transfer indicated that there were significant differences between conditions, Wilks'  $\lambda = 0.66$ ,  $F(6, 178) = 6.74$ ,  $p < .001$ . Regarding retention, the impact of condition was significant, as indicated by the ANCOVA,  $F(3, 90) = 9.43$ ,  $p < .001$ . Post-Hoc pairwise comparisons, based on Bonferroni's test, revealed that participants in the elaborative feedback condition outperformed those in the corrective feedback condition ( $p = .001$ ), the prompts condition ( $p = .001$ ), and the control condition ( $p < .001$ ), which did not differ from each other ( $p$ 's = 1). The size of these effects was large ( $d = 1.15$ ,  $d = 1.06$ , and  $d = 1.40$ , respectively). With regard to transfer, there was a significant effect of condition,  $F(3, 90) = 13.56$ ,  $p < .001$ . Participants in the elaborative condition outperformed those in the corrective feedback condition ( $p < .001$ ), the prompts condition ( $p < .001$ ), and the control condition ( $p < .001$ ), which did not differ from each other ( $p$ 's  $\sim 1$ ). The size of these effects were large ( $d = 1.43$ ,  $d = 1.23$ , and  $d = 2.23$ ).



Table 1: Performances of all conditions in all variables.

	Elaborative feedback		Corrective feedback		Prompts		Control	
	M	SD	M	SD	M	SD	M	SD
Prior Knowledge	2.44	2.20	2.10	1.62	1.96	1.39	1.38	1.76
Retention	11.63	4.16	7.35	3.22	7.30	4.04	6.08	3.76
Transfer	10.50	3.53	5.12	3.98	5.49	4.54	3.83	2.33

## Discussion

The results clearly revealed that elaborative feedback was more effective than corrective feedback and prompts, which were not more helpful than no support. This means that, for learners with little prior domain knowledge expected to gain a deep understanding of a complex topic, a great amount of support may be needed. Similarly, at least under these conditions (complex topic, low prior knowledge), prompts and corrective feedback may be not helpful enough. Both results emphasize the difficulty learners experience when having to self-regulate their learning.

It should be noted that participants were undergraduate students. It could be expected to find similar results with younger students, who probably have worse self-regulatory skills.

Prior studies used prompts and corrective feedback finding that they were effective (Azevedo et al., 2008; Lee et al., 2010), which differs from our results. One explanation is that these studies actually mixed different kinds of support within each condition, providing more support than that given here. In fact, Azevedo et al. not only prompted participants to self-regulate their learning but also set learning goals, asked learners questions to check their understanding, and, whenever they detected flaws in learners' understanding, directed learners to specific points in the material in order to solve the flaws. Similarly, Lee et al. not only provided feedback but also a set of prompts to highlight important sentences, summarize, and check and revise ongoing understanding.

A limitation of the experiment presented here was that we did not measured self-regulation directly but we inferred it from learning outcomes. Future research should examine self-regulation behaviour in some way (e.g., through think-aloud protocols) to ensure that the support devices affected it indeed and that learning outcomes result from such behaviour.

## References

- Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, 30, 87-111.
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research & Development*, 56, 45-72.
- Lee, H. W., Lim, K. Y., & Grabowski, B. L. (2010). Improving self-regulation, learning strategy use, and achievement with metacognitive feedback. *Education Technology Research & Development*, 58, 629-648.
- Mayer, R. E. (2009). *Multimedia learning* (2nd Edition). New York: Cambridge University Press.

# What Do Representations Say? – An Analysis of Students' Verbalizations

*Marion Geiger<sup>1</sup>*

*Markus Vogel<sup>2</sup>*

*Tina Seufert<sup>1</sup>*

<sup>1</sup>University of Ulm, Department of Psychology and Pedagogy  
Albert-Einstein-Allee 47 89069 Ulm, Germany

<sup>2</sup>University of Education Heidelberg, Institute of Mathematics and Computer Science  
Im Neuenheimer Feld 561, 69120 Heidelberg, Germany  
marion.geiger@uni-ulm.de

**Abstract.** Representations became more and more important over the last few years. Due to this students need skills to read off information out of a representation as well as to interpret different forms of representations. But studies show that students have problems to verbalize the information of representation. But where do these problems rely on? Concerning verbalization we can differentiate three levels: mathematical syntax, mathematical semantic and contextual semantic. We wanted to find out on which level students performed best and spontaneously. In our study 88 students took part and answered 18 tasks. Results show, that they spontaneously try to put information into a context of their everyday life but they performed better when they answered on level of mathematical syntax and only described what they see.

**Keywords:** Representation, verbalization

## Introduction

Representations, like graphs, tables or formulas, play an important role in school mathematics. Therefore students have to deal with them a lot. They often have to read information off a given representation and to transform it into another representation. For example, they see a graph of a linear function and have to produce the formula of this function. Besides these typical mathematical tasks, we have to deal with different representations in our everyday life: With circular or bar charts we illustrate for example results of election or with line charts someone can see the process of his performance during a computer game (e.g. money, produced cereals). Therefore, everybody needs certain skills to understand and interpret representations. But studies show that students have problems with transformation tasks (Ainsworth, 1999) and perform worst if they have to explain the message of a representation in their own words (Stradtman, 2010.)

Due to this, we wanted to have a closer look at processes of verbalization. Concerning these processes we can identify three different levels: mathematical syntax, mathematical semantic and contextual semantic. If someone verbalizes on the level of mathematical syntax he talks only about the appearance of the representation, for example the number of columns and rows of a table. A verbalization on the level of mathematical semantic goes beyond the description of the representation, because this level is characterized through the meaning of the symbols (Ruf & Gallin, 2005). But beside this is it possible to interpret a representation with regard to a special situation and to read of information which can't be seen at first view (Posner, 2003).

This can be illustrated with the aid of the following example: Imagine someone has to verbalize the information of the graph in figure 1.

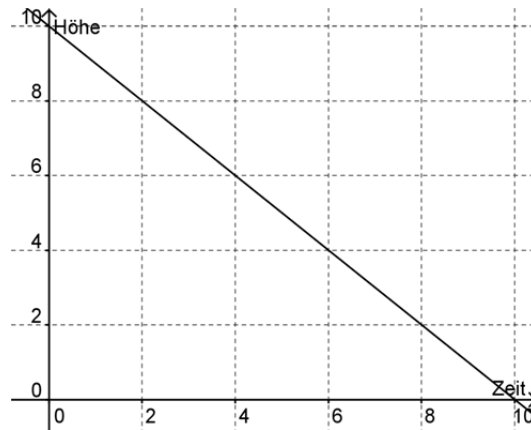


Figure 1. Example of a representation.

A possible verbalization on the level of mathematical syntax can be “*There is a coordinate system and one line, which rises.*”, while the answer on the level of mathematical semantic could be “*a linear function with y-intercept 10 and slope -1*”. And finally a verbalization on the level of contextual semantic could be read as follows: “*The candle was at the beginning 10cm high and decrease 1cm per hour*”.

In our study we analysed verbalizations in the domain of mathematics. The aim was to find out on which level students verbalize spontaneously and on which level the outcome of their verbalization is the best. Based on Stradtman (2010) we expected that students spontaneously verbalize on level of mathematical syntax and that verbalization on level of mathematical syntax would be best.

## Method

In order to test our hypotheses we developed 18 tasks in the domain of linear functions for all types of verbalization and all possible representations (graph, formula and table). Learners were given for example a formula and had to explain what they spontaneously think. To give students a chance to verbalize on level of contextual semantic axis of graphs or columns of tables were not labelled only by x and y but by words of everyday life, e.g. time or altitude. In other tasks students were asked by a special prompt to verbalize on one of mentioned levels.

88 nine-graders (age:  $M = 14.45$ ,  $SD = 1.58$ ; sex:  $m = 33$ ) took part in our study. In a within-subject design we analysed whether tasks with different prompts led to different results and what students tell spontaneously. As dependent variable we evaluated the produced verbalization with regard to the level of verbalization (mathematical syntax, mathematical semantic or contextual semantic) and content of accuracy.

## Results

The analysis of data revealed that most students verbalize spontaneously on level of contextual semantic if they had to explain what they think when they see a graph ( $\chi^2(2) = 75.89$ ,  $p < .001$ ) or a table ( $\chi^2(2) = 66.68$ ,  $p < .001$ ). But with regard to formula, we found no effect ( $\chi^2(2) = 3.65$ , n.s.). Students preferred no level of verbalization when they see a formula.

To find out on which level students performed best, we conducted an ANOVA with repeated measures. The ANOVA showed an main effect ( $F(2,74) = 12.93$ ,  $p < .001$ ,  $\eta^2 = .26$ ) for the level of verbalization. Therefore, we analysed contrasts and found that students performed better on level of mathematical syntax than on level of mathematical semantic ( $F(1,75) = 24.29$ ;  $p < .001$ ;  $\eta^2 = .25$ ) and on contextual semantic ( $F(1,75) = 24.29$ ;  $p < .001$ ;  $\eta^2 = .19$ ).

## Discussion

We found that students performed spontaneously on level of contextual semantic. They tried to relate the information of the representation to their everyday life. But as the results of our second analysis show they have enormous problems to verbalize on level of contextual semantic whereas they performed best on level of mathematical semantic. This shows, that students process representation only low level and didn't manage to put information in a context.

But this study has one limitation: We asked students to tell what they spontaneously think when they see a representation. And as mentioned above axes and columns were labelled by words of everyday and students answered on level of contextual semantic. We didn't know whether they answer on this level if columns and axes are labelled by  $x$  and  $y$ .

## Reference

- Ainsworth, S.E., (1999) A functional taxonomy of multiple representations. *Computers and Education*, 33(2/3), 131-152.
- Posner, R. (2003). Ebenen der Bildkompetenz. In K. Sachs-Hombach (Ed.), *Bildwissenschaft: Vol. 10. Was ist Bildkompetenz? Studien zur Bildwissenschaft* (1st ed.). Wiesbaden: Dt. Univ.-Verl.
- Ruf, U., & Gallin, P. (2005). Austausch unter Ungleichen: Grundzüge einer interaktiven und fächerübergreifenden Didaktik (3. Aufl.). Dialogisches Lernen in Sprache und Mathematik. Seelze-Velber: Kallmeyer.
- Stradtman, U. (2010). Analyse der Fähigkeiten zur Übersetzung zwischen verschiedenen Darstellungsformen in der Mathematik. Ulm.

# Studying and Executing Procedures: Do Pictures Facilitate Visuo-Spatial Reasoning During Learning?

*Erlijn van Genuchten<sup>1</sup>, Charlotte van Hooijdonk<sup>2</sup>, Katharina Scheiter<sup>1</sup>, Anne Schüler<sup>1</sup>*

<sup>1</sup> Knowledge Media Research Center  
Schleichstraße 6, 72072, Tübingen, Germany  
e.genuchten@iwm-kmrc.de

<sup>2</sup> VU University Amsterdam  
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands  
c.m.j.van.hooijdonk@vu.nl

**Abstract.** In this study we investigated whether pictures facilitate spatial information processing when studying procedural information. Participants studied first aid procedures, which differed in complexity and which were presented as text or as text and pictures. Working memory involvement during learning was gauged by measuring the interference between learning and the performance of a secondary spatial motor task. Participants performed the procedures immediately after learning and also after one week. As expected, participants learning with text and pictures performed the procedures more accurately than participants learning with text only. However, this difference between learning with text and pictures and text only reduced over time. Also, performing a secondary motor task interfered only with processing visuo-spatial information from text-only. Unexpectedly, this effect was the same for low and high complex tasks.

**Keywords:** multimedia, procedures, working memory, dual task, computational offloading

## Introduction

Research on multimedia learning has repeatedly shown that pictures aid learning (i.e., multimedia effect; Mayer, 2009). One explanation for this multimedia effect may be that pictures are especially suited to convey visuo-spatial information (Larkin & Simon, 1987), which makes them important when visuo-spatial reasoning is required for understanding the learning material. For example in procedural tasks, visuo-spatial information is important, because it comprises the spatial relationships between actions, which are intended to achieve a goal (Brunyé, Taylor, Rapp, & Spiro, 2006). Here, visualisations may provide immediate access to this information, making them computationally more efficient than text. In text, visuo-spatial information has to be interpreted (Larkin & Simon, 1987), which requires more computational effort in working memory (Scaife & Rogers, 1996). Therefore, it can be argued that – when visuo-spatial reasoning is required – the load on visuo-spatial working memory (i.e., visuo-spatial sketchpad [Baddeley, 1986]) is larger when processing text only than when processing text and pictures.

In this study, participants learned procedural tasks, differing in complexity, with text only or with text and pictures. Learning outcomes were measured immediately after learning and also after one week, and the dual-task approach was used to measure working memory involvement. First, we hypothesised that participants learning with text and pictures would perform procedural tasks better than participants learning with text only (i.e., multimedia effect). Second, we hypothesised that performing a spatial motor task interferes more with learning when studying text-only, because imagining visuo-spatial information requires more working memory processing capacity than extracting visuo-spatial information from pictures (cf. computational offloading [Scaife & Rogers, 1996]). Third, we hypothesised that the interference between the spatial motor task and learning is higher for high complex tasks, because more visuo-spatial reasoning is required to understand these tasks. Finally, as pictures facilitate visuo-spatial reasoning during learning, there is no reason to believe that the size of the multimedia effect changes over time.

## Method

### *Participants and design*

Eighty-eight students (74 female;  $M = 22.96$  years) participated. The experiment had a 2x2x2x2 mixed design, with picture presentation and dual task as between-subject variables, and task complexity and time of testing as within-subject variables. Depending on picture presentation and dual task condition, participants received 1) a text without dual task, 2) a text with a dual task, 3) a multimedia instruction without dual task, or 4) a multimedia instruction with a dual task. Participants studied two low and two high complex tasks and procedure performance accuracy was measured immediately after learning and also after one week.

### *Materials and procedure*

The procedural tasks were four first aid tasks, which were obtained from the Orange Cross manual (Henny, 2006). The procedures' complexity was defined by the number of steps (Robinson, 2001). Two low complex tasks contained 4 or 5 steps, whereas two high complex tasks contained 10 or 12 steps. In the multimedia instructions, two or three pictures were used per task. The dual task used to load the visuo-spatial sketchpad during learning involved tapping four pedals on the floor, which interferes with spatial reasoning in working memory (Miyake, Emerson, Padilla, & Ahn, 2004). During their performance of the first aid tasks, participants' actions were filmed from two angles.

Participants first familiarized themselves with the first aid puppet by performing a simple first aid task. Next, they filled out a demographic questionnaire. Participants who had to perform a dual task during learning practiced this task for one minute. Then, all participants studied the four learning tasks, which were presented in random order. Subsequently, they performed the tasks using a puppet and described the steps following or preceding a depicted step. One week later, participants performed the tasks in the same order. Performance accuracy was measured by the percentage of steps that was both performed correctly and in the correct order. To assess the inter-rater reliability, two raters coded 20% of all videos ( $\kappa = .65$ ). The remaining 80% was scored by a single rater only.

## Results

A 2x2x2x2 repeated-measures ANOVA was performed, with presentation format and dual task condition as between-subject variables, and task complexity and time of testing as within-subject variables. The dependent variable was the performance accuracy of the learnt procedures. Only effects that tested the hypotheses are reported.

The significant main effect of presentation format ( $\eta_p^2 = .27$ ), showed that participants learning with text and picture ( $M = .69$ ) performed the first aid tasks more accurately than participants learning with text only ( $M = .57$ ). The significant interaction between presentation format and dual task ( $\eta_p^2 = .07$ ), showed that performing a spatial motor task indeed interferes only with learning when studying text only (see Figure 1). Furthermore, the expected 3-way interaction between presentation format, dual task, and task complexity ( $\eta_p^2 = .02$ ) was non-significant, implying that the interference between the dual task and picture processing was the same for low and high complex tasks. Finally, the interaction between time of testing and presentation format ( $\eta_p^2 = .14$ ) was significant, indicating that the difference between participants learning with text and pictures, and learning with text-only reduced over time.

Table 1: Means for the interaction between presentation format and dual task.

	Dual task	Without dual task		With dual task	
		Text-only	Text-picture	Text-only	Text-picture
Time of testing	Presentation format	M	M	M	M
Immediately after learning	Low	.64	.79	.46	.77
	High	.61	.68	.44	.58
One week later	Low	.76	.76	.62	.77
	High	.58	.63	.43	.57

## Discussion

We investigated whether spatial reasoning is facilitated by pictures when studying procedural information. As hypothesised, there was a multimedia effect; however, the difference between students learning with text and pictures, and with text-only reduced over time. One explanation is that the performance of participants in the text-only conditions was improved after a week in one of the low complex tasks (immediately after learning:  $M = .69$ ; one week later:  $M = .79$ ), because the picture that was shown during the second post-test for this task helped participants to remember how to perform the procedure. Furthermore, performing a dual task interfered only with processing visuo-spatial information from text-only. Unexpectedly, the interference between the dual task and picture processing was the same for low and high complex tasks. One explanation could be that even though the number of steps is larger in high complex tasks, the level of difficulty of individual steps was identical between low and high complex tasks. When it is not required to integrate steps across time (which requires keeping all steps in working memory), the visuo-spatial sketchpad is unaffected by the number of steps. Together, these findings suggest that pictures in procedural tasks facilitate reading off visuo-spatial information and therewith the effortless interpretation of this information.

## References

- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology*, 20, 917-940.
- Henny, W. (2006). Orange Cross Manual, official manual for providing first aid (25th ed.). Utrecht: ThiemeMeulenhoff.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.
- Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, 115, 123-142.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.
- Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal Human-Computer Studies*, 45, 185-213.

# Picture-Word Learning by Spacing the Learning Sessions

*Nicole Goossens*

*Gino Camp*

*Peter Verkoeijen*

*Huib Tabbers*

*Rolf Zwaan*

Erasmus University Rotterdam

Burgemeester Oudlaan 50, Room T13-15, 3000 DR Rotterdam, The Netherlands

goossens@fsw.eur.nl

**Abstract.** In vocabulary learning pictures are commonly used. In this paper we investigated if spaced learning can enhance picture-word learning. The spacing effect is the phenomenon that distributing learning across time leads to better retention than massing it into one single study session. During three sessions, children of grade 1 were taught the meaning of eighteen words and their related pictures. In the massed condition, nine words were divided into three sets of three words, and each set was taught three times in one of three learning sessions. In the spaced condition, learning was distributed across the three sessions: All nine words were practiced once in each of the three learning sessions. At both the productive task and the receptive task after four or five days we observed that the meaning of spaced words was remembered better than the meaning of massed words.

**Keywords:** vocabulary learning, pictures, spacing effect

## Introduction

Vocabulary size is a powerful predictor of reading comprehension. In fact, researchers agree that between 90 and 95 percent of the words in a text need to be known to arrive at an adequate reading comprehension of the text (Hirsch, 2003). Therefore, it is important to stimulate vocabulary development already at an early age.

From research on pictorial and verbal cues we have learned that processing supportive information such as pictures enhances language learning. For example, Kellog and Howe (1971) found that foreign words which were learned with images or actual objects were remembered better than those without such additional information. One of the characteristics of good vocabulary instruction, listed by Blachowicz, Fisher, Ogle, and Watts-Taffe (2006), is that words have to be taught intentionally during multiple exposures in which definitional and contextual information is provided. Thus, repetition of unknown words is important and contextual information such as pictures can help to learn the words. However, not only the kind of repetition is important, but also how the repetitions are distributed over time. For example, words can be learned several times during one learning session (*massed learning*), but they can also be distributed over several learning sessions (*spaced learning*). The phenomenon that spaced repetition leads to better retention than massed repetition is commonly referred to as the spacing effect (for a review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). The question in this paper is whether spacing vocabulary learning sessions will lead to better retention of word meanings than massing these sessions.

The spacing effect has been demonstrated in several studies, in which children were learning pictures, words, or a combination of pictures and words (e.g. Rea & Modigliani, 1987; Toppino, 1993). However, those spacing experiments bear little resemblance to primary school vocabulary lessons. In the experiments the spacing intervals and the retention intervals were very short. The children who participated in these studies had to learn the words during the same learning session. The spacing intervals varied between 0 and 5 seconds or between 0 and 3 intervening items. These



intervals are in no way representative of learning intervals employed in a real vocabulary learning situation. In primary school vocabulary learning lessons, these intervals are much longer.

Further, most studies on the spacing effect with primary school children used free recall tests in which children had to recall as many words as they remembered (e.g. Toppino, 1993). In real vocabulary learning tests people use receptive tasks and productive tasks (e.g. Verhoeven & Vermeer, 2006). In receptive tasks children are asked to *name* an object (e.g. 'How do we call this object' by showing a picture of a coin) and in productive tasks children are asked to *explain* the meaning of a word (e.g. 'What is a coin'). It is still an open question whether the spacing effect generalizes to these intervals and these tests.

In this experiment we used long intervals and both tasks because they are representative of a real vocabulary learning situation. We expected the productive task to be more difficult than the receptive task because the children need more knowledge of a word when they have to explain it than when they have to name it. We hypothesized that we would find a spacing effect on both tasks. Because the spacing effect has not been tested with these tasks before we also explored if there would be an interaction between condition and task.

## **Method**

### *Participants*

A total number of 43 children from grade 1 participated in all sessions of the experiment. The mean age of these 43 children was 6.91 years ( $SD = 0.41$ ). The children were recruited from two classes from a medium-sized primary school, situated in Rotterdam. The children knew they participated in an experiment and their parents had given informed consent for participation.

### *Design & Materials*

In this experiment, we used a within-subjects design. We manipulated the distribution of words within learning sessions (massed learning versus spaced learning).

We selected eighteen concrete nouns and their definition from grade 2 learning materials. The words were first presented in a story told by the experimenter. Later, the words were presented with a picture and their definition (e.g. *A market vendor – Somebody who sells fruit on the market.*). This presentation of words was repeated in several learning sessions. The test consisted of two tasks and the children had to do both tasks, namely a productive task in which they had to define the word (e.g. *What is a market vendor?*) and a receptive task in which they had to name the picture (e.g. *How do we call the person on this picture?*).

### *Procedure*

The experiment took place in a classroom setting. At the start of the experiment, the children were told that they were going to learn some new words with the new vocabulary teacher (the experimenter) and that they would be given an unspecified test after four or five days. There were four learning sessions on four consecutive days and one test session four or five days after the final learning session.

Session 1 was an instruction session about the eighteen vocabulary words. In this session the experimenter told a story in which the words were included. Afterwards, all words were presented orally to the children in a PowerPoint presentation along with a picture and a definition. In Session 2, 3 and 4 the children listened to one oral explanation of nine words in the spaced condition and to three oral explanations of three words in the massed condition. All eighteen words were practiced three times during Session 2, 3 and 4. Within the session there was a short break in which the children did some filler tasks. Four or five days after the fourth session, the children received a test.

## Results

The mean percentage of retention on the productive and receptive task of massed and spaced words is shown in Table 1. We analysed the results by using a 2 x 2 ANOVA. There was a significant effect of condition,  $F(1,42) = 17.616$ ,  $p < .001$ ,  $\eta_p^2 = .295$ . There was no effect of task,  $F(1,42) = 3.266$ ,  $p = .078$ ,  $\eta_p^2 = .072$ . There was no interaction effect between condition and task ( $F < 1$ ). This means that for both tasks, performance was better in the spaced condition than in the massed condition.

Table 1: Mean percentage of retention of massed and spaced words (with *SD* in parentheses).

	Massed	Spaced
Productive Task	67.44 % (18.05)	75.97 % (16.61)
Receptive Task	69.51 % (20.15)	79.59 % (14.73)

## Discussion

This study shows that spacing words across several learning sessions is better than massing them into one learning session when children are learning vocabulary in an educational context. Rea & Modigliani (1987) and Toppino (1993) found already a spacing effect for children in picture-word learning by using very short spacing intervals and by using free recall tasks. In our experiment we found a spacing effect by using relevant spacing intervals and by using relevant tasks which are used in an educational context.

A number of questions still remain. For example, there is the question if spacing could reduce the number of repetitions needed to retain a word. Nation (1990) argues that the required number of exposures to retain a word varies from 5 to 16, but it is not clear if spacing could reduce this required number of exposures. Another question is if spacing would also enhance performance on more complex vocabulary. Donovan and Radosevich (1999) found in their meta-analysis of the spacing effect that there was a negative relationship between the complexity of the task and the effect size of the spacing effect. In our experiment, we used concrete nouns which did not require much effort in processing. Further research is needed to investigate the spacing effect using more complex vocabulary tasks.

## References

- Blachowicz, C. L., Fisher, P. J., Ogle, D., & Watts-Taffe, S. (2006). Vocabulary: Questions from the classroom. *Reading Research Quarterly*, 41 (4), 524-530.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354-380.
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect. *Journal of Applied Psychology*, 84, 795-805.
- Hirsch, E. D. (2003). Reading Comprehension Requires Knowledge – of Words and the World: Scientific Insights into the Fourth-Grade Slump and the Nation's Stagnant Comprehension Scores. *American Educator*, 27 (1), 10-29.
- Kellog, G., & Howe, M. J. (1971). Using words and pictures in foreign language learning. *Alberta Journal of Educational Research*, 17 (2), 87-94.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Rea, C. P., & Modigliani, V. (1987). The spacing effect in 4- to 9-year old children. *Memory & Cognition*, 15, 436-443.
- Toppino, T. C. (1993). The spacing effect in preschool children's free recall of pictures and words. *Bulletin of the Psychonomic Society*, 31 (1), 27-30.
- Verhoeven, L. & Vermeer, A. (2006). *Verantwoording Taaltoets Alle Kinderen (TAK)*. Arnhem: Cito

# What's the Point of Bulletpoints? The use of Text in PowerPoint Lectures

*Madeline Hallewell*

Learning Sciences Research Institute, University of Nottingham  
Exchange Building, Jubilee Campus, Wollaton Road  
Nottingham, NG8 1BB, United Kingdom  
ttxmh18@nottingham.ac.uk

**Abstract.** Undergraduate lectures are often accompanied by slideshows containing text outlines made available to students prior to the lecture in order that they may print them out as hand-outs upon which to take notes. This practice has been criticized for resulting in repetitive lecturing. The extent to which such text based slides are actually repeated by the lecturers during their lectures is examined through analysing transcripts of 21 lectures, and from focus group interviews with students, and copies of these students' notes from 10 of these lectures. It was found that although students expect their lecturers to use the text as a scaffold for the lecture; the extent to which this scaffold is obeyed by the lecturer varies considerably. This inconsistency can result in difficulties for students following the lecture, or their deliberately ignoring the lecturers' speech. The paper highlights the predicament of lecturers in planning their communications using text slides.

**Keywords:** Lectures, PowerPoint, Text- Talk integration

## Introduction

The tendency for displaying text in lecture slideshows has emerged from an embracing of slideware technologies in Higher Education, such that lecture slideshows now predominantly take the form of text outlines which the lecturer talks through point- by- point during the lecture (Olliges, Mahfood, Seminary, & Tamashiro, 2005). Moreover, with the spread of virtual learning environments (VLEs), lecturers have the means to not only show the slides during the lecture, they can disseminate them to students beforehand. The practice is popular amongst both lecturers and students as it affords the ability for lecturers to prepare an outline of the topics to be covered during the lecture which students can a) print out beforehand and use as a framework to take notes of the extra information given during the lecture, and b) revise from before assessments. However this "triple delivery" mode of instruction in which the student sees the text on screen, hears it spoken by their lecturer and also reads it on their hand-out has been criticized for being repetitive, (Maxwell, 2007), which can result in redundant information and therefore increased "cognitive load" for those receiving it (Mayer, 2001). Further, the inclusion of text specifically in slideshows has resulted in widespread criticisms including that it results in the use of stunted sentences (i.e. bulletpoints), thus stunting cognition (Tufte, 2006). More worryingly though, it has been suggested that slides have a speech suppression effect on students as they perceive the slides to contain the most important information, and as such unrelated speech is mostly ignored (Wecker, 2012).

It has been suggested that the practice might result in the expectation within students that the lecturer will take them through the slides with their speech, expanding and explaining the topics along the way (Adams, 2006). This expectation might explain the speech suppression effect, as students selectively listen out for anything that relates to the slide text. However, there is little currently known about the impact the slideshow outline practice has on the way in which lecturers communicate during lectures. The aim for the research was to consider the text based lecture slideshow in relation to the way in which lecturers communicate during lectures, thus the extent to which lecturers really do repeat their slides with their speech is questioned here.

## The Research

The research used a mixed methods approach to consider 2 questions; 1) When text is included in a slide lecture, to what extent is it integrated and repeated by the lecturer's speech? And 2) how does the extent to which the speech repeats (or in fact doesn't repeat) the structure of the slide text impact on the students' experience of slide lectures.

Initially, a corpus of video recordings of 11 introductory lectures given on Attachment Theory in undergraduate Psychology courses across the UK during the academic year 2009/10 was collected. These lectures were transcribed to present the sections of lecturers' speech alongside the corresponding slide that was displayed at the time. A second corpus of 10 further Psychology lecture recordings during the following academic year was then collected and transcribed in the same way. Additionally for the secondary corpus, focus group interviews were carried out with a selection of the students in attendance, using the lecture video recordings and slide hand-outs as stimuli for discussion. Copies of the participating students' notes were also collected.

## Methods of Analysis

Repetition of the slide text can be carried out linguistically by the lecturer through the means of "secondary pointing procedures" (Knoblauch, 2008) in which the speech mirrors or "parallels", to differing extents, the content of the slide. The transcripts then were analysed using a content analysis of the slide text to identify what meanings *could be* pointed to or repeated, and a conversation analysis of the speech utilizing this "paralleling" framework to identify where the speech actually repeated (or otherwise referenced) the slide text. Then, by reading the slide text from top to bottom and left to right, the slide text was turned into a string of letters (e.g. A, B, C, D etc) to represent the expected pattern of repetition by the lecturer. Using the order in which the slide text was actually repeated by the speech, strings were produced for the pattern of repetition of text by the speech accompanying each slide (for example A, D, C). Comparing these strings using the Levenshtein edit distance algorithm enabled a statistical expression of the extent to which the text of each slide was integrated as expected by the lecturer's speech.

A qualitative thematic analysis of the interviews afforded an investigation of the students' reactions to the speech-slide relationship. Thus their reflections were investigated in order to consider the impacts of the lecturers' speech- slide relationship on the student's experience in the lecture. Further, the students' notes were examined to consider what information had been noted and to identify the origin of this information (i.e. from the speech only, the slide only or a combination of the two).

## Findings

In analysing the lecture transcripts the Levenshtein edit distance statistic revealed that lecturers rarely matched their slide structure perfectly. Further the extent to which slide text representations were actually repeated as expected with the lecturer's speech differs both within and between lecture, resulting in often unpredictable patterns of integration. Scaling the scores such that 0 indicated no match and 1 indicated a perfect match, the maximum mean integration score received by the lecturers was 0.89 and the minimum 0.66. A one way ANOVA was used to test for differences between the lecturers in the extent to which their actual patterns matched their expected patterns using the individual slides as the population. Lecturers differed significantly in the similarity of the speech to the slide,  $F(10, 364) = 3.801$ ,  $p < 0.001$ .

Interview data indicate that the majority of the student participants agreed that the role of the PowerPoint slides was to provide some form of structure for their note taking. Further it was clear that the participants agreed having the "main points" available on a handout was beneficial to both note taking and for revision purposes. This has resulted in the expectation amongst these students that the lecturer will address the slide information in a systematic point- by- point manner in the lecture.

Moreover, analysis of the students' notes revealed that the majority of notes taken consist of information from the speech that serves to make sense of the text on the slide, for instance definitions of words or phrases, or adding words to make the bulletpoints into full sentences. During the focus groups, students reasoned that speech which did not immediately appear to be directly related to the slide text was "irrelevant" and so not deemed noteworthy, which was reflected in their lack of notes on this additional information.

### **Discussion and Implications**

Although criticized in much of the literature, the literal repetition of slide text by speech in PowerPoint presentations was not found in the lectures. It is suggested that uncertainty and unpredictability in lecturers' repetition of slides might have negative impacts on how students assimilate the information from the speech with the information from the slides. Student data highlighted a common understanding of the form of communication expected during slide lectures; that the slide text will be repeated and explained by the speech, and any other information is unimportant. Thus if lecturers give information not represented on the slide, or are cryptic in their integration of the text such that students perceive that the lecturer is violating the scaffold of their slides, students, on finding it difficult to match the speech to the corresponding text, may deem the speech to be "irrelevant" causing them to "switch off" and potentially miss crucial information during note taking.

This paper highlights the predicament of lecturers when considering giving a slide lecture that contains text. Lecturers are expected to provide the "main points" of the lecture in this format in order for students to take notes on them, but their doing so impacts on the way in which they are expected to communicate during the lecture and what students deem important or not. It seems that lecturers must negotiate a difficult balance when communicating using text representations in their lectures, thus it is suggested that lecturers reconsider their use of text in their lecture slides. Further research is clearly needed to discover the extent to which learning processes are either helped or hindered by differing patterns of repetition.

### **References**

- Adams, C. (2006). PowerPoint, habits of mind, and classroom culture. *Journal of Curriculum Studies*, 38(4), 389-411.
- Knoblauch, H. (2008). The performance of knowledge: Pointing and knowledge in Powerpoint presentations. *Cultural sociology*, 2(1), 75-97.
- Maxwell, A. (2007). Ban the bullet-point! Content-based PowerPoint for historians. *The History Teacher*, 41(1), 39-54.
- Mayer, R., E. (2001). *Multi Media Learning*. Cambridge: Cambridge University Press.
- Olliges, R., Mahfood, S., Seminary, K. G., & Tamashiro, R. (2005). From Talking At You to Talking With You: Reshaping PowerPoint for Interactive Learning. In M. Thirunarayanan, O & A. Perez-Prado (Eds.), *Integrating technology in higher education* (pp. 63). Maryland: University Press of America.
- Tufte, E., R. (2006). *The Cognitive Style of PowerPoint: Pitching Out Corrupts Within* (2 ed.). Cheshire, CT: Graphics Press LLC.
- Wecker, C. (2012). Slide presentations as speech suppressors: When and why learners miss oral information. *Computers & Education*, 59(2), 260-273. doi: 10.1016/j.compedu.2012.01.013

# Learning with Dynamic Visualizations: The Role of the Human Mirror Neuron System

*Birgit Imhof<sup>1</sup>, Ann-Christine Ehlis<sup>2</sup>, Peter Gerjets<sup>1</sup>*

<sup>1</sup> Knowledge Media Research Center, Schleichstr. 6, 72076 Tübingen, Germany

<sup>2</sup> Department of Psychiatry & Psychotherapy, University of Tübingen, Calwerstr. 14, 72076 Tübingen, Germany  
b.imhof@iwm-kmrc.de, ann-christine.ehlis@med.uni-tuebingen.de, p.gerjets@iwm-kmrc.de

**Abstract.** This study investigates whether viewing gestures facilitate learning about non-human movements and whether correspondence between gesture and to-be-learned movement leads to better results than non-correspondence. Functional near-infrared-spectroscopy was used to address whether gestures activate the human mirror-neuron-system (MNS), whether this activation transfers to subsequent viewings of non-human movements, and whether the activation of the MNS mediates the facilitation of learning. During learning participants viewed triples of visualizations for four movements (1. animation depicting a fish movement, 2. video with human gestures, 3. replay of the initial animation). Preliminary results showed that viewing gestures and viewing the replay of the animation (only in the case of correspondence) leads to a higher MNS activation. Furthermore, on a descriptive level, results indicated that higher MNS activation is associated with better learning performance. Thus, stimulating the MNS by means of gesture-based interventions might be an adequate instructional strategy to enhance learning with dynamic visualizations.

**Keywords:** Learning about movements, dynamic visualizations, human mirror-neuron-system, gestures, functional near-infrared-spectroscopy

## Learning with Dynamic Visualizations

Many contents in the Natural Sciences comprise the understanding of changes over time and space. Dynamic visualizations can easily depict such changes and they may be particularly suited for instructional purposes if these changes do not occur in discrete steps, but rather involve continuous aspects. In order to use dynamic visualizations effectively and to exploit their potential for learning it is crucial to understand when and for whom they are beneficial. Until now, research on dynamic visualizations has yielded rather heterogeneous results: Not only design factors and individual learner characteristics, but also context factors, such as, the knowledge domain or task requirements, influence their effectiveness (e.g., Höffler & Leutner, 2007; Tversky, Morrison, & Bétrancourt, 2002).

A novel hypothesis that has recently received considerable attention claims that the stimulation of specific areas in the brain is beneficial for learning about continuous processes with dynamic visualizations (e.g., Van Gog, Paas, Marcus, Ayres, & Sweller, 2009). These areas in the brain (i.e., the so-called mirror-neuron-system [MNS]) are used to understand and imitate the actions of other persons. They are typically activated by human movements, but may be more generally used to also represent other biological or even non-biological movements (if these movements can be anthropomorphized by the observer, cf., Engel, Burke, Fiehler, Bien, & Rösler, 2008). In the domain of learning about continuous processes, one effective instructional strategy to activate the MNS might be to show learners not only the to-be-learned processes, but also gesture-based interventions displaying the to-be-learned dynamics in order to trigger an anthropomorphized encoding. It has already been shown that the production of such gestures during learning is beneficial for acquiring knowledge (e.g., Cook & Goldin-Meadow, 2006; Scheiter, Arndt, Imhof, & Ainsworth, 2012).

Until now, there is no direct test of the assumption that learners' ability to recruit their MNS during processing dynamic visualizations may influence the effectiveness of the visualizations. Moreover, it still has not been investigated whether MNS activation can be induced by gesture-based interventions and then transferred to non-human movements because of mapping processes. Therefore, this study tested by using neurophysiological methods (i.e., functional near-infrared-spectroscopy [fNIRS]) whether the MNS is activated during viewing gestures and whether this activation transfers to subsequent viewings of non-human movements because learners map the human movements to the non-human movements. Moreover, this study investigated whether viewing gestures that correspond

to non-human movements facilitate learning about these movements better than non-corresponding gestures. Furthermore, this study tested whether the activation of the MNS mediates the facilitation of learning. We hypothesize that viewing corresponding gestures facilitates learning more than viewing non-corresponding gestures. Moreover, we hypothesize that human gestures activate the MNS, whereas this activation only transfers to subsequent non-human movements in the case of correspondence between gestures and non-human movements, because this helps learners to anthropomorphize the non-human biological movements. Furthermore, we hypothesize that learners differ with regard to recruiting the MNS for processing and that higher MNS activation is associated with better learning outcomes than lower MNS activation.

## **Methods and Results**

The present study with forty university students is currently under experimentation. Learners had to learn to discriminate four different fish movements. For each fish movement they saw a sequence of visualizations consisting of triples: Firstly, the learners saw an animation of a specific locomotion pattern. Secondly, they saw a video showing a gesture in form of human hand movements. These gestures either did or did not correspond to the locomotion patterns (experimental manipulation; between-subjects design). Thirdly, the learners saw the initial animation of the fish movement again. With this design, we were able to investigate: 1) whether biological movements activate the MNS (activation during the first view of the fish animation); 2) whether videos of corresponding and non-corresponding human gestures activate the MNS (activation during viewing the video); and 3) whether the biological movements activate the MNS more strongly after the learners have seen the gestures and whether this depends on the correspondence between the gesture and the to-be-learned movement (activation during the second view of the fish animation). An expert regarding fish movements performed the gestures. For the corresponding gestures this expert was instructed to display the respective movements as clearly as possible with his hands, whereas for the non-corresponding gestures the expert was instructed to perform gestures with his hands that had nothing to do with fish movements (i.e., waving, pointing, drumming, and circulating the forearms around each other). Each visualization was depicted for 30 s and followed by pauses of 30 s (black screen) between all visualizations. In the pauses, the activations of the brain areas of interest are supposed to decay to the baseline level before the next visualization was displayed. Following the learning phase, participants performed a locomotion pattern identification test that measured learning outcomes. This test consisted of 24 underwater videos of real fish (in blocks of 30 s) for which the learners had to classify the depicted movement. During learning and testing fNIRS measurements were conducted (ETG-4000, Hitachi). The probe set (2x22 channel array) was placed over fronto-temporo-parietal regions (centered at T3/T4 and C3/C4). Changes of absorbed near-infrared light were transformed into relative concentration changes of oxygenated (O<sub>2</sub>Hb) and deoxygenated haemoglobin (HHb). Local increases of O<sub>2</sub>Hb as well as decreases of HHb are indicators of cortical activity (Obrig & Villringer, 2003).

First results showed that there was a difference in the activation of the MNS between the three visualizations in a triple (animation for the first time before gestures, videos with gestures, animation for the second time after gestures). In particular, the activation of the areas that are associated with the human MNS (Brodmann areas 40, 44, and 45, Fogassi & Ferrari, 2010) was higher when learners viewed the gestures than when they viewed a fish movement animation for the first time. Additionally and more interestingly, the results also indicate that the activation of the MNS is higher when learners viewed the fish movement animation for the second time after they saw the corresponding gesture, but not after they saw the non-corresponding gesture. Because the number of subjects already studied is rather small, we could not test the expected correlation between the amount of activity in the MNS and the learning outcomes on a statistical level yet. Nevertheless, the results that will be fully available at the time of the SIG2 meeting already point in the expected direction on a descriptive level.

## Discussion

This study tested whether the human MNS can be activated for non-human biological movements by using gesture-based interventions intended to stimulate anthropomorphization. It has to be noted that the anthropomorphization is stimulated by an external video and is not accomplished by the learners on their own. Nevertheless, the results indicated that looking at human hand gestures activated the human MNS. Moreover, first results pointed in the direction that this MNS activation transferred to a subsequent viewing of an animation of the to-be-learned fish movements in the case that the human gestures corresponded to the to-be-learned contents. Therefore, learners seem to be able to activate the MNS during processing non-human biological movements if they have once been demonstrated a connection between these movements and movements of the human body. If it turns out that there is statistical evidence that higher MNS activation is associated with better learning outcomes, the strategy used in this study could be a promising device in different instructional domains and settings that involve learning about continuous movements and processes. Thus, it could turn out that gesture-based instructions that support anthropomorphization are a suitable method to enhance learning about processes with dynamic visualizations. However, in order to assess the real beneficial impact of the human gestures to the MNS activation further research has to investigate whether other interventions, such as for example viewing mechanical dynamic scenes, leads to the same results.

## References

- Cook, S. M., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, 7, 211 - 232.
- Engel, A., Burke, M., Fiehler, K., Bien, S., & Rösler, F. (2008). What activates the human mirror neuron system during observation of artificial movements: bottom-up visual features or top-down intentions? *Neuropsychologia*, 46, 2033-2042.
- Fogassi, L., & Ferrari, P. F. (2011). Mirror systems. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 22-38. doi: 10.1002/wcs.89
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738.
- Obrig, H., & Villringer, A. (2003). Beyond the visible – Imaging the human brain with light. *Journal of Cerebral Blood Flow & Metabolism*, 23, 1-18.
- Scheiter, K., Arndt, J., Imhof, B., & Ainsworth, S. E. (2012). *Move Like a Fish: Do Gestures Aid Learning from Photographs and Videos?* Manuscript in preparation.
- Tversky, B., Morrison, J., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.
- Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror-neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review*, 21, 21-30.



# The Impact of Discrepancies across Web Pages on High-School Students' Trustworthiness Evaluations

*Yvonne Kammerer, Peter Gerjets*  
Knowledge Media Research Center  
Schleichstrasse 6. 72076 Tuebingen, Germany  
y.kammerer@iwm-kmrc.de, p.gerjets@iwm-kmrc.de

**Abstract.** This study investigated how discrepancies across Web pages affect high-school students' source evaluations when reading multiple Web pages on a complex health-related topic. In line with the Documents Model framework and previous research on single-text comprehension, the present results indicate that content-based discrepancies across Web pages stimulate greater source evaluation, such that students mentioned more sources in their summaries and judged a Web page with obvious commercial interests as less trustworthy than when the Web pages contained consistent information. In the latter case, students did not seem to consider potential commercial biases of information sources in their trustworthiness evaluations.

**Keywords:** World Wide Web, online reading, multiple documents, source evaluation, discrepancies

## Introduction

With the rise of the World Wide Web, the online reading of multiple documents (i.e., Web pages) to learn about a particular topic has become a common reading situation, for instance, when searching for health information. However, on the Web not only official institutions, but also companies and laypeople provide information. These information providers might have other goals than just providing neutral facts, such as selling or promoting a particular product or exchanging experiences or stating opinions. Hence, it is important that Web users critically evaluate the trustworthiness of information sources that they retrieve from the Web and weight and interpret a Web page's content in light of its source characteristics.

Based on the Documents Model framework (e.g., Rouet, 2006) we hypothesized that readers would be more likely to evaluate and encode information about Web sources (e.g., document type, suspected motives) and draw relationships between sources and content ("who said what") when different Web pages contain discrepant information (cf. Bråten, Britt, Strømsø & Rouet, 2011). To the best of our knowledge, the presence of content-based discrepancies between multiple documents, however, has not yet been directly manipulated in any experimental study. Though, research on single-text comprehension provides first evidence that readers pay more attention to sources and also mention more sources in summaries when reading short news stories that convey discrepant information from two sources as compared to consistent information (Braasch, Rouet, Vibert, & Britt, 2012; Rouet, Britt, Caroux, Nivet, & Le Bigot, 2009). The goal of the present study was to extend these findings to reading contexts involving multiple online-documents and to test the hypothesis that readers would refer more to the sources (i.e., the different Web pages) in their summaries and would rate biased, commercial pages as less trustworthy when the Web pages encountered online provided discrepant information as compared to consistent information.

## Method

Forty-eight eleventh grade students of a German private high school (mean age: 18.2 years; 75% female) were given the task to seek information about the effectiveness of L-Carnitine to enhance athletic performance, in order to give informed advice to a fictitious friend. Participants' prior knowledge on the subject matter was generally low.

To complete the experimental task, participants were provided with a Google mock-up showing a search results page that contained five search results linked to five fabricated Web pages: Web page #1 was a shop, #2 a forum, #3 the Website of the German Nutrition Society (thereafter referred to as institutional page), #4 a commercial information portal, and #5 the personal blog of a European champion in bodybuilding. Each Web page contained about 400 words and reported studies that had investigated the effects of L-Carnitine supplementation on athletic performance. Participants were randomly assigned to two groups: Half of the participants received discrepant information, such that four Web pages reported that studies had shown the effectiveness of L-Carnitine for improving athletic performance, whereas the institutional page referred to studies that had demonstrated its ineffectiveness. The other half received consistent information, with all five Web pages reporting research that had shown the effectiveness of L-Carnitine. Participants were given 10 minutes to read the Web pages, but also could stop reading earlier. Participants were allowed to take notes during reading. Subsequently, participants were asked to give a recommendation whether or not L-Carnitine is effective for enhancing athletic performance and to summarize what they had found out about the issue, for which they could use their notes. Finally, participants were presented all five Web pages again in random order and had to rate the trustworthiness of each Web page on a 5-point rating scale.

## Results

Participants on average read 3.54 Web pages ( $SD = 1.40$ ) for which they spent 8.84 minutes ( $SD = 1.57$ ), without any differences between groups (both  $ps > .80$ ). Three participants of the discrepant information condition who had not visited the institutional page were excluded from all further analyses, as they could not have detected any discrepancies. After the search, 33.33% of the students in the discrepant information condition judged L-Carnitine to be ineffective as compared to 0% in the consistent information condition ( $\chi^2(1, N = 45) = 9.47, p < .01$ ). 57.1% at least mentioned the discrepancy in their summary. Furthermore, participants who were provided with discrepant information referred more to the information sources (i.e., the type or name of the Web pages) in their summaries than participants with consistent information,  $t(22.43) = -2.23, p = .04$ . However, also in the discrepant information condition source references still were very seldom, with 0.62 source references per student as compared to 0.08 source references in the consistent information condition. Finally, with regard to students' trustworthiness evaluations, an ANOVA with information discrepancy (discrepant vs. consistent) used as between-subjects factor and Web page type (the five different Web pages) used as within-subjects factor showed a significant interaction between information discrepancy and Web page type  $F(4, 172) = 3.25, p = .01$ . Posthoc tests revealed that students rated the shop significantly less trustworthy when provided with discrepant information than with consistent information ( $p < .001$ ). For the trustworthiness ratings of the four other Web pages no differences were found between the two groups (all  $ps > .17$ ). For means and standard deviations see Table 1. When comparing students' trustworthiness ratings across the five Web pages, results showed that in the discrepant information condition students rated the institutional page significantly more trustworthy than the shop, the forum, and the blog (all  $ps < .05$ ). The commercial information portal, however, was also rated significantly more trustworthy than the shop ( $p < .001$ ). In contrast, in the consistent information condition the trustworthiness ratings between the shop and the institutional page did not differ significantly ( $p = .16$ ). Only the two Web 2.0 sites (forum and blog) were rated as significantly less trustworthy than the institutional page ( $p = .001$  and  $p < .001$ , respectively) and the commercial information portal (both  $ps = .04$ ).

Table 1: Trustworthiness ratings for the five Web pages as a function of information discrepancy.

Web page	Discrepant information	Consistent information
Shop	1.90 (0.22)	3.25 (0.20)
Forum	2.33 (0.24)	2.58 (0.23)
Institutional page	3.95 (0.19)	4.00 (0.18)
Commercial information portal	3.14 (0.21)	3.54 (0.20)
Blog	2.57 (0.22)	2.63 (0.21)

## Conclusions

In line with the Documents Model framework and previous research on single-text comprehension, the present results indicate that content-based discrepancies across Web pages stimulate greater source evaluation, such that high-school students mentioned more sources in their summaries and judged the shop as less trustworthy than when the Web pages contained consistent information. Interestingly, in both groups the Web 2.0 sources were judged as less trustworthy, indicating students' skepticism regarding these type of sources. However, the results are also alarming: first, in the discrepant information condition source references still were very seldom and two-thirds of the students believed in the effectiveness of L-Carnitine to enhance athletic performance. Second, whereas discrepancies led students rate the shop as rather untrustworthy, the commercial information portal that also promoted a commercial L-Carnitine product apparently was not identified as such. Moreover, when only consistent information was provided, the shop was judged as trustworthy as the institutional page of the German Nutrition Society. To conclude, when only retrieving Web pages with consistent information, high-school students do not seem to consider potential biases of information sources in their trustworthiness evaluations. A search engine that provides Web pages with discrepant information on a particular issue among the first few search results (i.e., those that are predominantly selected by searchers), in contrast, might be a way to increase students' source evaluations on the Web.

## References

- Braasch, J., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, 40, 450-465.
- Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Towards an integrated model. *Educational Psychologist*, 46, 48-70.
- Rouet, J.F. (2006). The skills of document use: From text comprehension to Web-based learning. Mahwah, NJ: Erlbaum.
- Rouet, J.-F., Britt, M.A., Caroux, L., Nivet, C., & Le Bigot, L. (2009). *The influence of story consistency and reading context on the construction of documents models from multiple sources*. Paper presented at the biennial conference of the European Association for Research on Learning and Instruction. Amsterdam, The Netherlands.

# Does Construction of Diagrams Deepen Understanding by Raising Awareness of Insufficiency in Learning?

*Sachiko Kiyokawa, Yotan Kura*

Department of Psychology, College of Humanities, Chubu University  
1200, Matsumoto-cho, Kasugai-shi, Aichi, 487-8501, Japan  
kiyo@isc.chubu.ac.jp and hw07036-1997@sti.chubu.ac.jp

*Yuri Uesaka*

Graduate School of Education, the University of Tokyo  
7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan  
y\_uesaka@p.u-tokyo.ac.jp

*Emmanuel Manalo*

Faculty of Science and Engineering, Waseda University  
3-4-1, Ohkubo Shinjuku-ku, Tokyo, 169-8555, Japan  
emmanuel.manalo@gmail.com

**Abstract.** Although previous studies have shown the beneficial effects of diagram construction and use on comprehension and problem solving, the mechanism that explains why constructing diagrams promotes understanding has not been clearly revealed. The present study examined the hypothesis that students become meta-cognitively aware of what they do not understand when they construct diagrams during learning of sentential text information, thus promoting deeper understanding. One hundred and sixteen undergraduates were assigned to one of three conditions: learning with diagrams, with words, and without any instruction (control). Participants were asked to learn a text about blood circulation by following instructions according to their condition, and a questionnaire assessing their perception of difficulty and understanding was subsequently administered. The results showed that perception of difficulty was highest among participants in the learning with diagrams condition compared to the other two conditions. In contrast, assessment of understanding was lowest in the learning with words condition. The results partially supported the hypothesis.

**Keywords:** Diagrams, text comprehension, monitoring

## Introduction

Diagrams are considered to be effective strategies for understanding and problem solving. For example, Butcher (2006) found that diagrams supported mental model development and comprehension of an expository text. Previous studies have also revealed the mechanism of how the facilitative effects of diagrams can be obtained when diagrams are provided: for example, Larkin and Simon (1987) explained that in diagrammatic representations related information is located in one place (which is not the case with sentential representations), and features like this make using diagrams more computationally efficient.

Although many studies involving diagrams have focused on providing diagrams, students' construction of diagrams facilitates greater effects on understanding and problem solving. For example, Stern, Aprea, and Ebner (2003) empirically demonstrated that students who actively construct linear graphs as reasoning tools while learning economics evidenced better transfer effects compared to a receptive condition in which learners received the diagrams to use rather than constructing their own. However, the mechanism of how the facilitative effects of diagrams can be obtained when participants are asked to draw diagrams by themselves has not been sufficiently examined.

One reason why constructing diagrams contributes to effective learning may be that the construction process helps students become meta-cognitively aware about what they do not understand while learning the material in question, and this awareness leads to deeper understanding via more careful processing of the information provided in text. The act of constructing diagrams from sentences requires students to create new representations in their mind, and this may help in the elaboration and re-organization of information that they already possess. Thus, using diagrams during

learning of sentential material could cause in students an internal perception of difficulty in understanding (as they may also become more aware of inefficiencies associated with their understanding). Additionally, learning with diagrams may contribute to a deeper understanding of information represented in the material. This study experimentally tested these hypotheses.

## **Method**

### *Participants*

One hundred and sixteen undergraduates from Chubu University participated in the experiment and received course credit following completion of the experimental session.

### *Materials*

An expository text about the human circulation system, written in Japanese and containing 866 script characters, was made from a biology text (Suzuki, Tasaki, & Yamamoto, 1999). Diagrams were excluded in order to investigate the effects of self-generated diagrams on text comprehension.

### *Conditions*

Participants were randomly assigned to one of three conditions: diagram, word, or control. The conditions were different in the strategies participants were directed to use in the learning session.

### *Procedure*

The experimental sessions were conducted with groups of up to 30 participants at a time. The text was presented on a page of a booklet given to the students. Firstly, the participants were given 10 minutes to silently read and comprehend the text. In the diagram condition, the participants were then required to draw diagrams on an A4 sized blank sheet of paper to summarize the content of the text. The participants in the word condition were asked to use a similar sheet of paper to summarize the text using words. The participants in the control condition were instructed to use their sheet of paper as they liked in order to comprehend the text. After the learning session, the participants in all conditions were required to answer questions including one that assessed their perception of difficulty in understanding. After this, a filler task was administered for 30 minutes, in which the participants were asked to generate two ideas based on 30 news articles. After this filler task, another assessment was conducted to evaluate the participants' degree of understanding of the material. Among the tasks administered during this session was a task to generate three keywords, which was scored for appropriateness and used in this study as the measure of the participants' degree of understanding.

## **Results**

### *Perception of Difficulty in Understanding*

A one-way ANOVA was conducted using the participants' perception of difficulty in understanding as the dependent variable and the learning condition as the independent variable. The result revealed that the main effect of the learning condition was significant ( $F(2, 113) = 7.33, p < .01$ ). It indicated that the perception of difficulty in the diagram condition was greater than the other conditions ( $HSD = 0.49, p < .05$ ).

### *Keyword Score (Degree of Understanding)*

A one-way ANOVA was conducted using the keyword score (that was used as the measure of understanding) as the dependent variable and the learning condition as the independent variable. The result revealed that the main effect of the learning condition was significant ( $F(2, 113) = 6.20, p < .01$ ). It indicated that the understanding in the word condition was lower than the other two conditions ( $HSD = 0.88, p < .05$ ).

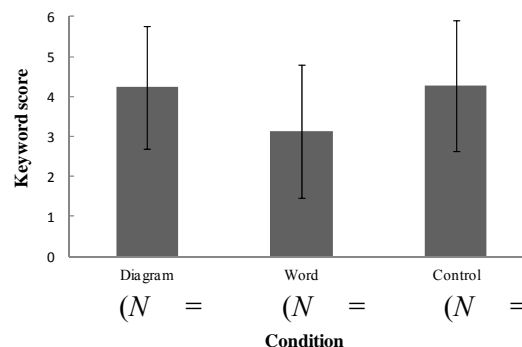
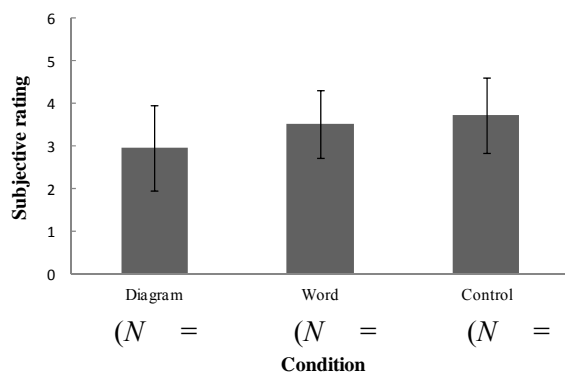


Figure 1. Mean and SD of perception of understanding Figure 2. Mean and SD of the keyword score

## Discussion

The result concerning perception of understanding demonstrated that the participants perceived greater difficulty when they were asked to construct diagrams during learning of the sentential material. To construct diagrams required them to change their mental representation to a different format, and this likely instigated a need to reconstruct and elaborate the information provided in text. As diagrammatic representations locate related information in one place (cf. Larkin & Simon, 1987), participants would have needed to consider relationships between concepts appearing in different places in the sentential text and to integrate those into new clusters in their diagrams. In this process, they were likely to have realized insufficiencies in their understanding – which may explain why those in the diagram condition perceived greater difficulty in understanding than those in the other two conditions.

However, the perception of greater difficulty during learning in the diagram condition does not mean that they did not understand: the subsequent score in understanding was higher in the diagram condition compared to the condition without diagrams (where they had to use only words). This result supports our idea that participants who were requested to draw diagrams during learning would process information provided in text more deeply (via reorganizing and elaborating) and this would help them better understand the material than participants who did not receive such an opportunity. However, the keywords score of participants in the diagram condition was equivalent to that of participants in the control condition – which does not support our hypothesis. A possible reason for this may be that the time allowed for learning the text (10 minutes) might not have been adequate: despite becoming aware of insufficiencies in their understanding, those in the diagram condition may not have had adequate time to sufficiently reconstruct their understanding of the information provided. The provision of more time to learn and the use of other measurements of understanding (apart from the keywords score) would need to be investigated in future research.

## References

- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98, 182–197.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.
- Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction*, 13, 191–203.
- Suzuki, T., Tasaki, K., & Yamamoto, T. (1999). *Daigaku katei no seirigaku (8th Ed.)* Tokyo: Nankodo.

# The Effects of Visualization Forms on Usability and Learning Outcomes – Dynamic Videos versus Static Pictures

*Alexandra König*

*Mona Stadler*

*Melina Klepsch*

*Tina Seufert*

University of Ulm, Department of Psychology and Pedagogy  
Albert-Einstein-Allee 47 89069 Ulm, Germany  
alexandra.koenig@uni-ulm.de

**Abstract.** In previous research, much debate has centered around whether animations or pictures facilitate a higher level of learning. The combination of the concept of usability with this research establishes an additional aspect, hardly discussed in previous literature. The aim of the present study was to determine the association of visualization forms to learning outcomes and usability. Participants viewed static pictures with written texts and dynamic videos with spoken texts. After viewing each video and picture, subjects completed usability questionnaires as well as knowledge and comprehensive questions. Dynamic videos with spoken texts showed higher usability while static pictures with written text showed higher learning comprehension. The results could be seen as additional evidence for a higher usability of dynamic videos with spoken texts and indicate their use for motion practices.

**Keywords:** Visualization Forms, Videos, Dynamic, Pictures, Static, Usability, Learning Outcomes

## Introduction

The concept of usability is originally a part of the human-computer-interaction research and the major indicators of interest should be the needs and aims of users. Based on the literature of Mangold, Vorderer and Bente (2004), a product of high usability should be characterized by high utility, simple handling and satisfactory for the user. According to Mangold et al. (2004) there are certain criteria of usability, which could be categorized in performance-related (effectiveness and efficiency) and non performance-related criteria (satisfaction, acceptance, favor, fun, entertainment and fascination). In the development of our usability questionnaire, we were guided by two of the most reliable usability questionnaires for the evaluation of software and websites: The Software Usability Measurement Inventory (SUMI) and The Website Analysis and MeasureMent Inventory (WAMMI).

Research concerning the comparison of learning content presentation as static pictures with written text or as dynamic videos with spoken text provides inconsistent findings. On the one hand, Höffler and Leutner (2007) claim obvious advantages of animations concerning learning outcomes with an effect of animations on learning outcomes of  $d = .76$  whereas Tversky, Morrison and Bétrancourt (2002) declare themselves strictly against a use of animations in learning contents.

The combination of the concept of usability with the comparison of visualization forms is an additional aspect which was hardly discussed in the previous literature so far. In the current study, we were interested in whether the way of presenting is related to the usability of information given in motion practices. According to the research of Höffler and Leutner (2007) it was assumed that dynamic videos accompanied by spoken texts showed a higher usability than static pictures accompanied by written texts in the context of teaching motion practices and procedural knowledge. Moreover, we examined the coherence of the form of presentation with learning outcomes.

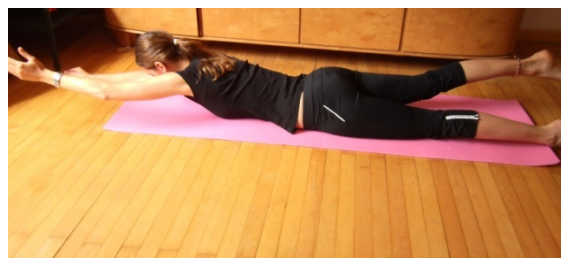
Based on Höffler and Leutner (2007) as well as on Mayer and Moreno (2002), we expected higher learning outcomes by presenting exercises as dynamic videos with spoken text than static pictures with written text.

## Method

In order to test our hypotheses, we developed a usability questionnaire and several recall- and comprehension questions concerning the presented content. The usability questionnaire consisted of four different scales. The first scale “effectiveness” (picture:  $\alpha = .81$ ; video:  $\alpha = .96$ ) should measure in which extent a certain target was reached, whereas the second scale “efficiency” (picture:  $\alpha = .65$ ; video:  $\alpha = .81$ ) specified the ratio of costs and extent of target achievement. Furthermore there were two more scales for the “way of presenting” (picture:  $\alpha = .69$ ; video:  $\alpha = .80$ ) and for the “expectation conformity” (picture:  $\alpha = .88$ ; video:  $\alpha = .96$ ). These four scales were finally summed up to the scale “usability” (picture:  $\alpha = .90$ ; video:  $\alpha = .89$ ).

The learning outcome questionnaire differentiated between the two forms of learning outcomes “recall” and “comprehension” and consisted of four different scales plus the scale “prior knowledge”. The first scale “recall video” and the second scale “comprehension video” were summed up to the scale “learning outcome video”. Likewise, the scale “recall picture” and the scale “comprehension picture” were summed up to the scale “learning outcome picture”. The scales had low alpha scores. However, we didn’t expect internal consistency to be high because of the broad variety of aspects we asked for.

In a within-subjects design, twenty-five learners participated in our standardized online study, in which all subjects ran through both conditions (the video as well as the picture condition) and answered the same questions concerning the presented motion practices. First, the participants were asked to score the usability of the shown visualization forms whereas they had to learn the information of the given motion practices in the second part. Therefore, in the picture condition the subjects were requested to look accurately at the static picture as well as to read the written explanatory text below the picture. In the video condition, their task was to concentrate on dynamic videos with spoken text. The videos and pictures contained information about the same four types of motion practices to advance the abdominal muscular system, the back muscles or the core muscles (example in figure 1), whereas the difficulty level of all motion practices was equal.



*Figure 1.* Example of a motion practice

Two separate ANOVAs with repeated measures with the way of presenting as independent factor were used to analyse usability and learning outcomes as dependent variables.

## Results

The analysis of the given data with an ANOVA for usability revealed an expected main effect for the visualization form: The usability score was significantly better for the videos ( $F(1, 24) = 8.58, p = .007, \eta^2 = .26$ ). The visualization as dynamic videos reached an average score of 4.41 ( $SD = .56$ ) in usability, whereas the visualization as static pictures only scored with an average of 4.20 ( $SD = .61$ ) on usability.

The second ANOVA for learning outcomes indicated no significant main effect of the way of presentation in the expected direction ( $F(1, 21) = .68, p = .42, \eta^2 = .03$ ). However, a main effect of presentation mode on comprehension in non-expected direction could be observed ( $F(1, 21) = 4.37, p =$



.05,  $\eta^2 = .17$ ). The visualization as dynamic videos reached an average score of .70 ( $SD = .28$ ) in comprehension, whereas the visualization as static pictures scored with an average of .84 ( $SD = .16$ ) on comprehension. Concerning the learning outcome, a marginal significant main effect could be shown, however in non-expected direction ( $F(1, 21) = 4.07, p = .06, \eta^2 = .16$ ).

## Discussion

To summarize the results, dynamic videos accompanied by spoken texts showed a higher usability than static pictures accompanied by written texts, but did not turn out to induce higher learning outcomes. According to Höffler and Leutner (2007), there could be shown that the usability of learning material on procedural-motor knowledge can be successfully supported by dynamic visualizations. But contrary to Höffler and Leutner (2007), static pictures accompanied by written texts evoke a significantly better comprehension performance as well as marginal significant higher learning outcomes. Thus, as indicated by Tversky, Morrison and Bétrancourt (2002), animations do not show advantages concerning learning outcomes in comparison to static pictures. One explanation for those findings is a possible cognitive overload for the learner's, provoked by the fast sequence of pictures in dynamic videos. Moreover, those results could be due to a higher distractibility of learners caused by an animated presentation by focusing on irrelevant details instead of the mere learning aspects (Höffler, & Leutner, 2007).

Further research needs to investigate whether these findings could be replicated in laboratory settings with a between-subjects-design with counterbalanced motion tasks each presented static and dynamically. Moreover, we should include and vary the second factor, modality of text presentation and improve the scales for learning outcomes. An interesting question for further research is also to include learner's prior knowledge. In any case, these findings are an additional evidence for the presented approach: The combination of the concept of usability with the comparison of visualization forms.

## References

- Hoeffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Media Learning and Instruction, 17*, 722-738.
- Mangold, R., Vorderer, P., & Bente, G. (2004). *Lehrbuch der Medienpsychologie*. Goettingen: Hogrefe.
- Mayer, R. E., & Moreno, R. (2002). Animation as an Aid to Multimedia Learning. *Educational Psychology Review, 14*, 87 - 99.
- Tversky, B., Morrison, J.-B., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human Computer Studies, 57*, 247-262.

# Biology Teachers' Professional Knowledge of Diagrams

*Lena von Kotzebue, Claudia Nerdel*

Technische Universität München, TUM School of Education, Fachdidaktik Life Sciences  
Schellingstraße 33, 80799 München, Germany  
lena.kotzebue@tum.de & claudia.nerdel@tum.de

**Abstract.** The analysis and construction of scientific graphs is an important attribute of the quality of scientific instruction. A current literature review identified graph-specific components of the teachers' professional knowledge based on two aspects: knowledge of graphs and models of teachers' professional action. Prospective biology teachers should have knowledge of these rules for the construction and interpretation of diagrams mentioned above. In addition to the identification of these components, we will develop subject-specific tests to be able to measure and operationalize these components. The first results of the pilot study will be presented at the EARLI SIG 2.

**Keywords:** learning with graphs, science education, teachers' professional knowledge

## Instruction

The analysis or construction of scientific graphs is an essential skill required in a school or scientific context and is becoming increasingly important in today's media age. Scientific graphs, however, are not intuitively comprehensible representations (Dreyfus & Eisenberg, 1990). Therefore quite a few students can neither understand these graphs properly nor use them effectively (Baumert, Stanat, Demmrich, 2001). In order to overcome these difficulties and to benefit from the advantages of scientific graphs, the interpretation and construction of these representations should be taught in scientific instruction. The teachers will be responsible for this knowledge transfer. Accordingly, teachers have to understand scientific graphs and they have to know how to use them in scientific instruction to get the benefit out of these representations. Due to this fact both the content knowledge (CK) of scientific graphs and their pedagogical content knowledge (PCK) play an essential role for prospective teachers. A study by McElvany and others (2009) found, however, that teachers often fail to correctly evaluate the difficulty of the tasks as well as the students' achievements when these use instructional images as learning media.

## Theoretical Background

Professional knowledge is an influential variable for successful teaching (Borowski, Neuhaus, Tepner, Wirth & Fischer, 2010). This knowledge refers to all theory-based parts of knowledge that teachers acquire during their training and the lessons they give at school (Clandinin & Connelly, 1995). The professional knowledge of a teacher is typically divided into three domains: content knowledge (CK), pedagogical content knowledge (PCK) and pedagogical knowledge (PK) (Schulman, 1987). In the following, it will be dealt with subject-specific components of the domains CK and PCK. Both are examined in this study.

### *Content Knowledge:*

a) Expert knowledge concerning the contents: profound background knowledge on the topics discussed in class is essential for the correct interpretation of the scientific graphs.

b) Methodological knowledge: the appropriate analysis or construction of scientific graphs refers to the three components of diagram competence as defined by Lachmayer, Nerdel and Prechtel (2007): the extraction of information from and the construction of scientific graphs as well as the integration of information derived from texts and diagrams.

### *Pedagogical Content Knowledge:*

a) Knowledge of the pupils' cognition: the teachers' knowledge about the pupils' difficulties or their misconceptions concerning the topic; furthermore, their ability to recognize and correct the pupils' methodological problems with scientific graphs.

b) Knowledge of the correct use of assignments: the choice of tasks with text and scientific graphs to structure the instruction. The teachers' knowledge of the cognitional potential of the tasks is essential for the appropriate use of scientific graphs in the lessons.

c) Knowledge of subject-related instructional teaching methods: aims, contents and methods should be developed as a coherent approach which ensures clear evidence of the interaction between the components.

### **Research Questions and Hypotheses**

The question arises if (prospective) biology teachers already possess such theory-based professional knowledge of how to use such graphs in their lessons more effectively. We assume that (prospective) biology teachers have a profound content knowledge which they have acquired during their studies and teaching training.

We suppose, however, that they have insufficient pedagogical content knowledge as the appropriate use of diagrams is not on the agenda of the teaching training at university and in schools. Thus teachers are not taught how to deal with conceptual change, i.e. they do not know how to reason about students' unobservable mental states that these use to predict and explain others' actions. Frequently, the teachers' diagnostic competence is not well-developed. This presumption is in line with a study by McElvany et al. (2009).

### **Research Design and Methodology**

In the pilot studies tests are constructed to analyse the factors of teachers' professional knowledge with regard to scientific graphs in biology mentioned above. The first pilot study assessed the content knowledge (CK) of photosynthesis and graphs, e.g. a subject-specific test of photosynthesis was developed to analyse the knowledge of scientific diagrams. The test is divided into three parts. It measures the knowledge of the topic photosynthesis (content knowledge - topic, CKT), methodological knowledge of using diagrams without the topic (content knowledge – method, CKM) and at last the use of diagrams in the aspects construction of scientific graphs and the integration of texts and graphs relating to photosynthesis (CKD). In this first pilot study 217 students participated.

A CFA model (estimator: WLSMV) was executed to validate these components of the theoretical model and to examine their relationships. The model provides five latent factors: CK (4 Items) and CKM (7 Items) as predictors for CKD. Furthermore, CKD was divided into three latent factors (1) construction of graph, information is provided by a text dealing with photosynthesis (CKD-C; 6 Items); (2) interpretation of graph by integrating the information of text and graph (CKD-II, 3 Items); (3) construction of graph by integrating the information of text and graph (CKD-CI, 3 Items). Correlations between CKD-C, CKD-II and CKD-CI were tested, too. The data were analysed by Mplus 6.11.

In a second pilot study the three components of the PCK will be analysed. At the moment the test is carried out with 200 students. Both parts of the test will be revised evidence-based and used in the main study to achieve a common model for all components mentioned above.

### **First Results and Discussion**

The first study tested the students' content knowledge (CK) on scientific diagrams. The influence of knowledge of photosynthesis (CKT) and methodological knowledge of graphs without a topic (CKM)

on the construction of scientific graphs and the integration of texts and graphs of the topic photosynthesis was measured. Figure 1 shows the correlations between the latent variables in detail.

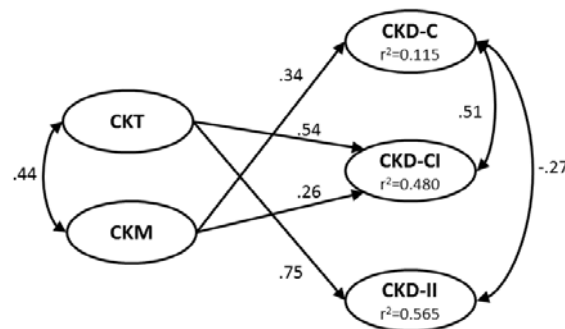


Figure 1. Correlations between the latent variables

The latent variables *construction* and *integration* of the model by Lachmayer et al. (2007) were confirmed. In our model integration was divided into two parts, namely CKD-CI and CKD-II. The knowledge of the topic photosynthesis predicts the integration of texts and graphs. The regression coefficient of content knowledge (CKT) is larger if the task requests a text (CKD-II) instead of a diagram (CKD-CI) as its solution. Furthermore, methodological knowledge of diagrams (CKM) predicts the construction of scientific graphs (CKD-C) as well as the integration of texts and graphs if the task requests a further diagram as its solution (CKD-CI). The regression coefficient for the prediction on integration is smaller because elements of the diagram that have to be constructed are already shown in the task. If the whole graph has to be constructed, students have to take the framing of the graph (axis, scales, legend etc.) into account, too. These findings are in line with Lachmayer (2008). Details of the CFA model and further results for PCK will be presented at EARLI SIG 2.

## References

- Dreyfus, T., & Eisenberg, T. (1990). On difficulties with diagrams: Theoretical issues. In G. Booker, P. Cobb, & T. N. De Mendicuti (Eds.), *Proceedings of the 14th Annual Conference of the International Group for the Psychology of Mathematics Education* (pp. 27–36). Oaxtepec: PME.
- Baumert, J., Stanat, P., & Demmrich, A. (2001). PISA 2000. Untersuchungsgegenstand, Grundlagen und Durchführung der Studie. In D. PISA-Konsortium (Ed.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 15–68). Opladen: Leske u. Budrich.
- Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., & Fischer, H. E. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN) – Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167–175.
- Clandinin, D. J., & Connelly, F. M. (1995). *Teachers' professional knowledge landscapes*. New York: Teachers College Press.
- Lachmayer, S., Nerdel, C., & Precht, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 161–180.
- Lachmayer, S. (2008). Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht. Unveröffentlichte Dissertation: Christian-Albrechts-Universität zu Kiel.
- McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W. et al. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für pädagogische Psychologie*, 23(3-4), 223–235.
- Schulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57, 1–22.

# Learning Art History on Multi-Touch-Tables: Metaphorical Meaning of Interaction Gestures matters.

Julia Kranz, Birgit Imhof, Stefan Schwan, Barbara Kaup, Peter Gerjets  
Knowledge Media Research Center, Schleichstr. 6, 72076 Tübingen, Germany  
j.kranz / b.imhof / s.schwan / p.gerjets@iwm-kmrc.de, barbara.kaup@uni-tuebingen.de

**Abstract:** This study investigated the influence of interaction gestures on a multi-touch-table (MTT) with regard to category learning. During learning participants learned to categorize Renaissance and Baroque paintings either by moving them over the display of the MTT (*dragging* gesture) or by pressing a marked field on the display representing an art epoch (*tapping* gesture). Moreover, the pressure necessary for interaction with the MTT was varied (low versus high). Subsequent to learning, participants had to categorize known and novel paintings. Two hypotheses were tested: Based on the idea of a *Gestural Conceptual Mapping*, it was hypothesized that the tapping gesture will be more beneficial for learning due its discrete character resembling the binary nature of categorization tasks. However, following an *Embodied Metaphor* approach, the dragging gesture should facilitate learning due to its similarity to the real world experience of sorting physical objects into piles. Results supported the embodied metaphor approach.

**Keywords:** reality based interaction, multi-touch-table, gestures, embodied interaction, art learning

## Introduction

Novel technologies like multi-touch-tables that enable learners to directly manipulate external representations with their body without using interaction device offer new options in designing learning environments. However, this leads to the issue of how these new options (like gesture-based interaction) should be deployed to support educational goals. Recently, growing evidence has been provided that gestures can affect thinking and learning (e.g., Goldin-Meadow & Beilock, 2010; Goldin-Meadow, Cook, & Mitchell, 2009). Segal (2011) coined the term *Gestural Conceptual Mapping* to describe the idea that learning can be supported by means of gestures that are congruent to mental processes necessary to achieve a learning objective. She observed children's performance in arithmetic tasks (involving discrete processes) and numerical estimation tasks (involving continuous processes) and found that discrete gestures supported arithmetic tasks whereas continuous gestures supported numerical estimation tasks. Based on the idea of a *Gestural Conceptual Mapping*, it can be hypothesized that other discrete processes, such as categorizing objects (biological, historical or other taxonomies) might also benefit from discrete gestures. For instance, tapping on a category might be a prototypical discrete gesture to support processes of categorization. On the contrary, one could argue that categorizing objects by means of a continuous gesture like moving these objects virtually to a place representing a category is much more similar to the action of sorting real physical objects, such as sorting papers on different piles or toys into different boxes. Therefore, it could be argued that a continuous gesture that forces learners to really move an object to a category is more suitable to support category learning because it uses an embodied metaphor for the mental process of categorization. The concept of an embodied metaphor is rooted in the Conceptual Metaphor Theory of Lakoff and Johnson (1980) as well as in the Embodiment Hypothesis claiming that thought is grounded in bodily experiences (Barsalou, 1999). Recent studies yielded embodied metaphors for different concepts: Boroditsky (2000), for example, provided evidence for the assumption that time is understood as space. Boot and Pecher (2010) showed that similarity is understood as spatial proximity. Most recently, Boot and Pecher (2011) demonstrated that the concept of categories is represented by means of a container schema. They presented pictures inside or outside a frame in order to activate the container schema and found that this influenced performance in congruence with the metaphorical mapping hypothesis (same category—inside bounded region; different category—not in same bounded region). In sum, there are two conflicting theoretical accounts that allow deriving predictions on the

effectiveness of MTT interaction gestures for category learning: Based on the idea of a Gestural Conceptual Mapping, learning should be more efficient when interaction gestures are congruent to the mental process necessary for learning. According to this approach, category learning should benefit from a discrete tap gesture that is aligned to the discrete character of the categorization process. On the other hand, moving objects to places that represent categories can be considered similar to the real world experience of sorting objects and might, therefore, serve as an embodied metaphor for the mental process of categorization. Based on this idea, category learning should be facilitated by using a continuous drag gesture.

## Methods

Sixty-five participants (mean age = 24.92;  $SD = 4.62$ ) with low prior domain knowledge individually participated in this study. The experiment consisted of a learning phase at the MTT and a testing phase using a laptop computer. The learning task was to categorize paintings according to their epoch (Renaissance or Baroque). In the learning phase at the MTT, participants were first presented with general information on Renaissance and Baroque paintings. Characteristic features of Renaissance and Baroque art were explained. Subsequently, subjects had to sort 14 randomized example paintings of each epoch into the right category. Paintings were entitled with the epoch they belonged to and subjects were instructed to reflect on the characteristics features introduced during sorting. During the testing phase, the 28 paintings known from learning phase and 28 novel paintings were presented in a randomized order and participants had to categorize them in a forced-choice task. We used a 2x2x2x2-design with the between factors interaction gesture (dragging a painting with the forefinger over the surface of the MTT into a category versus tapping on a place representing the category so that the painting would move automatically in the direction of the category as long as the category was touched) and necessary pressure for interaction (low versus high pressure needed to move objects) and the within-factors transfer distance of test items (known versus novel paintings) and epoch (Renaissance versus Baroque). Notice that for both interaction gestures the direction and length of the movements of paintings were the same, as was the visual information: In both groups, participants saw the painting moving to the category. To guarantee that learning took place simultaneously with the motoric activity, paintings were presented blurred and only became clear if the MTT was touched with enough pressure according to the respective gesture. As a result, the presentation time of a clear painting was held constant across conditions. The level of pressure necessary was additionally varied. As dependent variables response accuracies and reaction times in the test phase were used.

## Results

Subsequent to an item analysis, test paintings with a mean response accuracy below 50 % or above 95 % in any experimental group were excluded from the analysis, resulting in 47 test paintings. We submitted the response accuracies and the reaction times on valid responses for test paintings to two mixed-design ANOVAs, one based on participant variability ( $F1$ ), the other based on item variability ( $F2$ ). There was no speed-accuracy trade off. Accuracy and reaction times correlated negative in all groups. Response accuracies between the two gesture groups showed higher response accuracies in the *dragging* conditions than in the *tapping* conditions (*dragging*:  $M = 84.89\%$ ,  $SD = 6.52\%$ , *tapping*:  $M = 80.05\%$ ,  $SD = 10.20\%$ ;  $F1(1, 61) = 5.17$ ,  $p = .027$ ,  $\eta = .275$ ;  $F2(1, 43) = 22.27$ ,  $p < .001$ ). Response accuracies did not significantly differ between the two pressure groups in the  $F1$  ("low pressure":  $M = 83.82\%$ ,  $SD = 8.56\%$ ; "high pressure":  $M = 81.01\%$ ,  $SD = 9.09\%$ ,  $F1(1, 61) = 1.91$ ,  $p = .17$ , but in the  $F2$ ,  $F2(1, 43) = 4.51$ ,  $p = .04$ . With regard to reaction times, no effects of gesture or necessary pressure were found. Not surprisingly, there was an effect of transfer distance,  $F(1, 71) =$

56.09,  $p < .001$ . Paintings that were known from learning phase were categorized faster than novel paintings. Also, there was an interaction of transfer distance and necessary pressure,  $F(1, 71) = 4.52$ ,  $p = .037$ . Further, an effect of epoch was found,  $F(1, 71) = 20.11$ ,  $p < .001$ , as well as an interaction between epoch and transfer distance,  $F(1, 71) = 7.33$ ,  $p = .008$ . There were significant differences between known and novel paintings for both epochs but the effect of epoch was only significant for novel paintings.

## Discussion

First of all, the results provide evidence that the metaphorical meaning of interaction gestures matter for category learning on a MTT. This is an important finding, as there are only few studies showing a difference between gestures that transferred from a learning period to a test period (for example Goldin-Meadow et al., 2009). In particular, it was found that categorizing by dragging exemplars to a category is more beneficial for learning than tapping onto the right category. This result supports the embodied metaphor hypothesis postulating that the mental process of categorizing is grounded on the bodily experience of sorting physical objects into different places. Additionally, it provides further evidence for Boot and Pecher's (2011) assumption that the mental representation of the concept categories is represented by means of a container schema. Our finding of a learning advantage for the dragging gesture is not trivial, as the discreteness of the tapping gesture could have been beneficial, too. We did, however, not find any support for the Gestural Conceptual Mapping approach in this study, which requires further investigation. For instance, it is possible that categorization in our task was not as discrete as one might think. There might have been a "continuity of belonging to an epoch". Besides, there is also the possibility that both gestures facilitated learning because of different characteristics but that the dragging gesture was nevertheless superior. Future research has to further clarify these issues.

## References

- Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Boot, I., & Pecher, D. (2010). Similarity is closeness: metaphorical mapping in a conceptual task. *The Quarterly Journal of Experimental Psychology*, 63, 942-954.
- Boot, I., & Pecher, D. (2011). Representation of categories: Metaphorical use of the container schema. *Experimental Psychology*, 58, 162-170.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1-28.
- Goldin-Meadow, S. & Beilock, S. (2010). Action's influence on thought: the case of gesture. *Perspectives on Psychological Science*, 5 (6), 664-674.
- Goldin-Meadow, S., Cook, S. & Mitchell, Z. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20 (3), 267-272.
- Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.
- Segal, A. (2011). Do Gestural Interfaces Promote Thinking? Embodied Interaction: Congruent gestures and direct touch promote performance in math. New York: Columbia University.

# The Misinterpretation of Box Plots

Stephanie Lem<sup>1</sup>, Patrick Onghena<sup>2</sup>, Lieven Verschaffel<sup>1</sup>, Wim Van Dooren<sup>1</sup>

<sup>1</sup>Centre for Instructional Psychology and Technology, KU Leuven

Dekenstraat 2 – PO box 3773 – 3000 Leuven, Belgium

Stephanie.Lem@ppw.kuleuven.be

<sup>2</sup>Methodology of Educational Sciences Research Group, KU Leuven

Andreas Vesaliusstraat 2 – PO box 3762 – 3000 Leuven, Belgium

**Abstract.** Box plots have been increasingly present in academic, education and professional contexts. Recent studies, however, have shown that students have various difficulties interpreting these external representations for data distributions. In this study we focused on the misinterpretation of the area of the box plot as representing the number of observations in that interval. Using a dual process approach, we found that this misinterpretation is caused by heuristic reasoning based on the saliency of the area in box plots.

**Keywords:** external representations, box plots, dual process theories, graph design, saliency

## Theoretical and Empirical Background

Box plots are used increasingly in academic, education and professional contexts. However, studies have shown that students have great difficulty interpreting these representations of data distributions correctly (e.g., Lem, Onghena, Verschaffel, & Van Dooren, 2011, in press). Lem et al. (in press) suggested that students might reason intuitively or heuristically when interpreting box plots, considering the area of the box as representing the number or proportion of observations, while it represents density. This can be explained by Tversky's (1997) graph design principle on the use of space in graphs, which suggests that space should be used in a natural way as people tend to interpret for example area as an indicator of frequency or another relevant variable. Applied to box plots, we see that area does not represent frequency, but rather density.

According to dual process theories, two different types of reasoning can be employed when confronted with a task: heuristic and analytic reasoning. Heuristic processes are described as unconscious, automatic, fast, and undemanding; analytic processes as conscious, slow, deliberative, and effortful (Evans, 2008). Heuristic processes are frequently used, often leading to the correct solution. However, when heuristic reasoning leads to an incorrect solution, analytic reasoning is necessary. Then, the heuristic error needs to be detected and inhibited. According to the revised and extended heuristic-analytic model of Evans (2006), when an error has been detected in the heuristic reasoning process, (irrelevant) salient task features can still cause interference in the analytic stage of the reasoning process. Applied to box plots, the first impression is possibly that the larger part of the box represents the most observations. Even if one then starts to reason analytically, the influence of this first impression may not disappear, possibly leading to an incorrect response.

## Method

Participants were 118 first year university students at KU Leuven. All students had completed the same introductory statistics course several weeks before participation, covering box plots and descriptive statistics among various other topics. The students participated in return for course credit.

Participants solved 40 items in which they were given two box plots representing fictional exam scores of two classes. They had to determine in which of the classes more students scored higher than ten out of twenty. In the 16 *congruent* (C) items the heuristic response was the correct one, while in the 24 *incongruent* (IC) items the heuristic response was different from the correct response, making analytic reasoning necessary. Figure 1 provides an example of a C and an IC item.



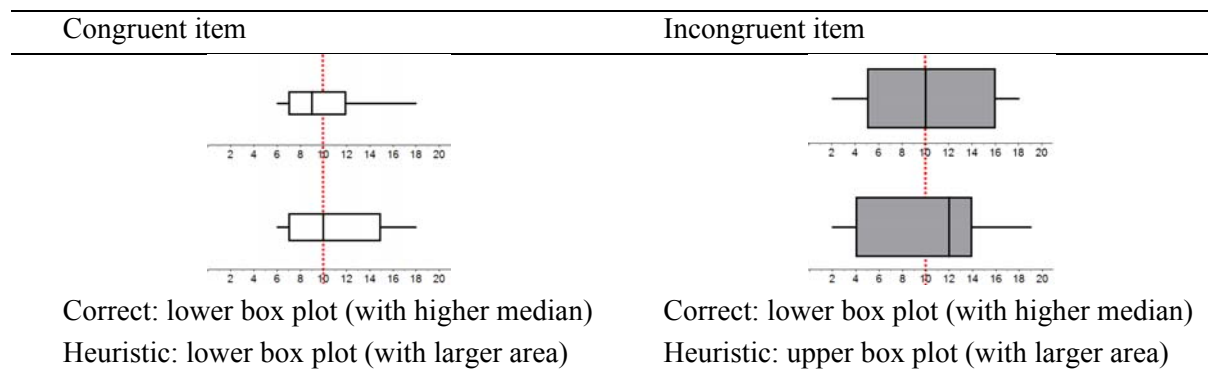


Figure 1. Example of a congruent and an incongruent item with correct and heuristic response

The same items were presented in a version in which the area of the box was more salient (e.g., higher and grey) and a version in which the area was less salient (e.g., lower and white). The saliency of the area of the box was varied randomly between subjects.

The participants were divided over three conditions. In the *heuristic condition*, the response time was restricted to 5000ms in order to prevent analytic processing to occur (or at least to complete). In the *analytic condition*, the use of analytic processes was encouraged: Participants had unlimited response time, were warned about the possible difficulties in interpreting graphs (with an example of an IC line graph item) and to carefully check their response before typing it in. The *control condition*, finally, did not have a limit on the response time, and did not get the warning of the analytic condition.

After the experiment, students were asked to answer ten questions about box plots. They had to name all elements of the box plot, explain what they represent, and they were asked to write down which percentages of the data were represented in the different parts of the box plot. This enabled us to assess whether incorrect responses in the experiment could be due to a lack of knowledge, rather than to the incorrect use of certain heuristics.

Dual process theory enabled us to make several predictions. First, we predicted higher accuracy for C items than for IC items, as in the IC items heuristic reasoning would not suffice. Second, we predicted most correct responses in the analytic condition and least correct responses in the heuristic condition, as the task conditions made it difficult to successfully employ analytic reasoning in the heuristic condition, while analytic reasoning would be activated by the warning given to the students and students would get time to complete this reasoning. Third, we predicted that salient items would be solved less well than non-salient items. Fourth, we predicted shorter reaction times to salient items than to non-salient ones, as the salient area of the box would lead to a fast conclusion based on this salient feature. Fifth, we predicted correct responses to C items to have faster reaction times than correct responses to IC items, as in IC items slower analytic reasoning is necessary to find the correct response. Finally, we predicted students' performance on the box plot knowledge test to be high.

## Results

A generalized linear mixed model analysis with accuracy as dependent variable was conducted to test the effect of congruency, saliency, and condition, as well as their interaction effects. We found a main effect of congruency, but not of saliency and condition. This main effect of congruency was in line with our first prediction: Accuracy was higher (89.7%) for C items than for IC items (61.7%). Surprisingly, our second prediction, concerning the effect of condition, could not be confirmed. We also found an interaction effect of saliency and congruency, congruency and condition, and of saliency, congruency, and condition. The interaction of saliency and congruency shows that saliency did not have an effect when solving C items, while it did when solving IC items. IC items were solved better (66.6%) with non-salient box plots than with salient box plots (59.4%). This follows our third

prediction. The interaction between condition and congruency can also be explained by dual process theories as the effect of congruency was largest in the heuristic condition, in which students had less time to overcome their heuristic response.

A generalized linear mixed model with the log reaction times of the correct responses as dependent variable was used to test the main effects of congruency, saliency, and condition, and the interaction effects between them. Three main effects were found. Salient items were solved faster (3014ms) than non salient items (3342ms), IC items were solved slower (3355ms) than C items (3002ms), and the heuristic condition was fastest (1934ms), while the analytic condition was slowest (4209ms). These main effects confirm our fourth and fifth prediction. The interaction effect of saliency and condition was also significant: The effect of saliency was largest in the control condition. Furthermore, the interaction of congruency and condition was significant: The effect of congruency on reaction time was largest in the control condition.

Finally, on average, students scored 94% on the box plot knowledge test, with 75% of all students answering all 10 questions correctly, confirming our final prediction. This shows that these students possessed the necessary knowledge to correctly solve the items in this experiment.

## Discussion

From this study we can conclude that the misinterpretation of box plots under study is not due to a lack of knowledge about box plots, but rather to inappropriate heuristic reasoning processes based on the salient feature area of the box plot. We have shown that Tversky's (1997) graph design principle on the use of space is an important factor that impedes the correct interpretation of box plots. Moreover, this principle seems to form the basis for heuristic reasoning processes. This suggests that there is, from a cognitive perspective, something fundamentally wrong with the design of box plots. Future research should investigate alternative designs of box plots and their effect on students' (mis)interpretations. Suggestions for alternative box plots can be found in Tufte (1983). Also studies with different groups of students or expert users of box plots would be interesting.

## References

- Evans, J. S. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, 13, 378-395.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgement and social cognition. *Annual Review of Psychology*, 59, 255-278.
- Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (2011). *External representations for data distributions: in search of cognitive fit*. Manuscript submitted for publication.
- Lem, S., Onghena, P., Verschaffel, L., & Van Dooren, W. (in press). On the misinterpretation of histograms and box plots. *Educational Psychology*.
- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire: Graphic Press.
- Tversky, B. (1997). Cognitive principles of graphic displays. In *Proceedings of the AAAI 1997 Fall symposium on reasoning with diagrammatic representations* (pp. 116–124).

# Secondary Students' Reading of Digital Visual Representations when using Physics Educational Simulations

*Víctor López Simó, Roser Pintó Casulleras*

CRECIM (Center of Research for Science and Mathematics Education)

Universitat Autònoma de Barcelona, Campus Bellaterra 08192, Edifici GL Despatx 304, Spain

victor.lopez@uab.cat, roser.pinto@uab.cat

**Abstract.** When secondary school students use simulations for physics learning, they have to read and understand the visual representations displayed in these educational tools. In this study, we analyse how students read three physics simulations and we identify the difficulties involved in this reading process, intending to relate them with their understanding of the three physics models which are expected students will construct supported by the use of each of the simulations. We have analysed these visual representations through different perspectives (social semiotics, psychology and science education) and we have carried out individual interviews in order to identify, categorize and analyse the most relevant visual features that bring up reading difficulties. Our final aim is to support science teachers to develop critical visual reading in the science class.

**Keywords:** Reading images, visual representations, educational simulations, physics education

## Rationale

Many secondary school physics teachers use free online available educational simulations (ES) for physics instruction because these educational tools include visual representations (VR) that can be very useful for enhancing students' comprehension of scientific models (Smetana & Bell, 2011). Teachers usually assume that these visual representations are always correctly read by students, and this leads teachers to think that the understanding of simulations' content (scientific models) is granted. However, previous researches on science education have shown that when students read scientific VR, they understanding may differ from the original meaning. It affects to the conceptual understanding of the scientific models which are expected to be learned by students (Ametller & Pintó, 2002; Cook, Wiebe, & Carter, 2008).

In this context, our research question is: Which reading difficulties appear when 14-16 years old students use some physics' educational simulations?

## Theoretical framework

Difficulties in the reading of VR have been studied and their results have been often stated as general principles. Social Semiotics has analysed the visual grammar of images, the compositional structure, the salience, the framing and the modality makers (Kress & van Leeuwen, 1996). Gestalt Theory analyses self-organizing tendencies of human brain and their influence in reading images (Winn, 1994). Multimedia learning theory has developed some general principles about learning with multimedia materials (Mayer, 2001). Studies on animated visual representations have analysed the effect of dynamism in the reading process (Lowe, 2003); and the theory about multiple external representations (MER) has analysed the design, functions and tasks of MER (Ainsworth, 2006).

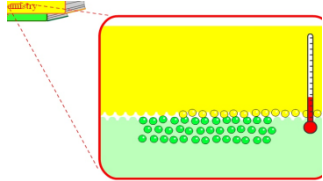
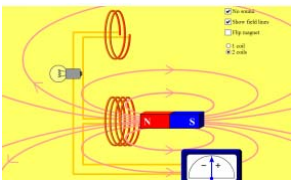
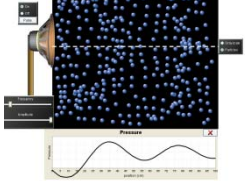
## Methodological design

The research methodology adopted in this study is qualitative since we are doing a **diagnosis** study that should inform us about the **existence** of reading difficulties but not a study about the **prevalence** of these difficulties.

First of all, we selected three simulations from an online library (<http://phet.colorado.edu/>). The selection process was done according to some curricular and educational criteria: content level, modality of their representations and minor importance of mathematics formalism. The three selected simulations were about heating by friction (A), electromagnetic induction (B) and propagation of

sound through mechanical waves (C). Later on, we developed a visual and content analysis of the three simulations in order to identify their relevant features, which were listed in a codification system.

Table 1: Three selected simulations

A: Heating by friction	B: Electromagnetic induction	C: Sound as a wave
		

We selected 40 students (14-16 years old) with some prior knowledge about these three scientific topics but not expert on them. We carried out a 20 minutes individual semi-structured interview per student, which were video-recorded, transcribed and analysed through the software *Atlas.ti*. During the analysis process, we selected, coded and classified pieces of the interview according the previous codification system based on the visual features of the simulations. At the end of this process, we generated a collection of difficulties faced by students during the reading process that we collapse into six independent **dimensions** (See Table 3).

Table 2: Example of analysis process for one specific analysis unit for simulation A

Relevant feature of simulation A	Quote	Researcher comment
Two surfaces in contact are represented, and students are asked to rub them. Then, some represented particles increase their vibration (microscopic representation of temperature rising) but some other particles disappear (mesoscopic representation of the surface erosion of any friction process).	<p>R: So, what can you observe when you rub the surfaces?</p> <p>S: I can observe the evaporation of some particles.</p> <p>R: Evaporation?</p> <p>S: Yes, some particles leave the screen.</p>	This student confuses erosion with evaporation. She has problems to distinguish between two different dynamic representation that have different modality because two different scientific scales (mesoscopic and microscopic).

## Results

The following table summarises the categories that allows us to answer the research question about which reading difficulties appear when students use the three physics' educational simulations:

Table 3: Summary of the most relevant difficulties in reading each simulation.

Dimensions	Reading difficulty	Examples from simulation A, B or C
D1. Compositional structure of the VR and framing of its visual elements	Not connecting the different parts of the compositional structure	Lack of recognition of the interdependence of the elements of an electric circuit (B)
	Not identifying one of the parts of the compositional structure	Interpreting the ground black square as a box (C)
D2. Salience of the visual elements of the VR	Over-attention to surface features	Interpreting the behaviour of the system in terms of the colour of the magnet (B)
	Over-attention to the text	Trying to explain the visual information in terms of the embedded text (A)
	Over-attention to the decorative elements	Interpreting the decorative shine of atoms as the atom's nucleus (A)
	Confusing different scientific	Confusing corpuscular and continuous

D3. Modality and visual value of the visual elements of the VR	models	models of matter (A)
	Confusing different scientific scales	Confusing meso, micro and macro scales (A)
	Confusing extensive and intensive ranges	Confusing atoms representation with atoms existence (A)
	Confusing different visual values	Conceptualizing magnet-field lines as real objects (B)
D4. Dynamism of the visual elements of the VR and	Confusing between position and movement of visual elements	<i>Interpreting the behaviour of the system in terms of position instead of in terms of movement (B)</i>
	Non-perceiving the non-linear changes	Lack of perception of the asymptotic rate of decrease of the thermometer (A)
D5. Interaction and relationship between multiple representations	Non-connecting complementary information	Lack of connection between particles representation and wave front representation (C)
	Non-discriminating different information presented at time	Confusing particles vibration and particles leaving (A)
D6. Distortion of the visual elements of the VR	Distorted perception of the shapes	Considering that the central hole of the coil implies that the circuit is opened (B)
	Distorted perception of the movement	Visual confusion between frequency of vibration and velocity of propagation (C)

## Conclusions

A wide range of typologies of reading difficulties have been identified, and each difficulty is associated to one or more visual features. There has been evidenced that science's specificities (such as models, scales and the nature of scientific representations) play an important role in the reading process. Thus, general statements about difficulties in reading images have to be reviewed from the content point of view.

Furthermore, although these difficulties come from specific simulations around three scientific topics, it has been possible to infer categories that appear in different situations, and for this reason they could be extrapolated to other physics' simulations. For example, the reading difficulty that we had classified as "*Confusing different scientific scales*" can potentially appear in other situations where students use a simulation that include representations of different scales. In the future, we expect to be able to formulate more general patterns.

Finally, once again we realise the important role of the critic reading of images that teachers should develop when their students use scientific simulations, scaffolding the images' reading process.

## References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198.
- Amettler, J., & Pintó, R. (2002). Students' reading of innovative images of Energy at secondary school level. *International Journal of Science Education*, 24(3), 285-312.
- Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, 92(5), 848-867. Wiley Subscription Services, Inc., A Wiley Company.
- Kress, G., & van Leeuwen, T. (1996). *Reading Images. The Grammar of Visual Design*. New York: Routledge.
- Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13(2), 157-176.
- Mayer, R. (2001). *Multimedia Learning*. New York: Cambridge University Press.
- Smetana, L. K., & Bell, R. L. (2011). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education*, DOI: 10.10.
- Winn, W. D. (1994). Contributions of perceptual and cognitive processes to the comprehension of graphics. . In W. Schnotz & Kulhavy (Eds.), *Comprehension of graphics. Advances in psychology* (Vol. 108, pp. 3-28). Amsterdam: Elsevier Science B.V.

# Is Making Written Syntheses an Aid to the Comprehension of Documentary Sources?

*Eva Lordán*  
eva.lordan@ub.edu

*Mariana Miras*  
mariana.miras@ub.edu

*Esther Nadal*  
esthernadal@ub.edu

*Sandra Espino*  
sespino@ub.edu

*Marta Minguela*  
martaminguela@ub.edu

*Cristina Luna*  
cristinaluna@ub.edu  
University of Barcelona

Pg. Vall d'Hebrón, 171. 08035, Barcelona, Spain

**Abstract.** This study was aimed at exploring the processes involved and the results attained by undergraduates in understanding information from multiple sources on two different tasks: reading/reading and writing a synthesis. The participants were 161 undergraduates –females and males- in Psychology. All of them were asked to read three texts about intelligence, and to answer a prior knowledge questionnaire –before reading- and a comprehension test –after reading-. The experiment was designed based on two different conditions: task (reading/reading&synthesis) and media (performing the task on paper or computer). Three ANOVAs were conducted with task and media as independent variables and global, superficial and deep comprehension as dependent variables. No significant differences were found between tasks. Significant differences were found on deep comprehension for media condition in favor of the students who did the task on paper.

**Keywords:** reading, written synthesis, reading comprehension, higher education

Understanding and integrating information from various sources, identifying the relations between them (contrasting positions, complementarity, causality, chronology, etc.) is a cognitively demanding skill (Mateos & Solé, 2009; Nelson Spivey, 1997). As opposed to what is required in reading a single source, a reader faced with several documents needs to find –and often create- a common thread enabling him or her to organise information from texts that may differ in important respects (structure, contents, length, density, genre, etc.) (Boscolo, Arfé & Quarisa, 2007; Segev-Miller, 2007). It is commonly thought that the intrinsic complexity of dealing with multiple sources can be mitigated by the use of writing: note-taking, making summaries and outlines, etc. Various studies have found that producing a synthesis that integrates different items of information enables people not only to retain, but also to elaborate and organise the information they have read in greater depth. Consequently, producing a synthesis would be expected to lead to a more complete and deeper understanding of texts than producing other types of text or just reading. Also, most students are faced with using ITC to perform reading and writing tasks, therefore there is a special interest in comparing the effect of both media –paper&pencil and computer- in the reading and writing processes and results students follow and obtain. Several studies have found that, when students are asked to read extended, difficult reading passages –as is the case of this study, those who perform the task on paper tend to obtain better results than those who perform on computer (Murphy, Long, Holleran & Esterly, 2003; Paek, 2005).

The aim of the study reported here was to examine the processes employed and the results attained by university students in understanding information from three texts on two different tasks: a) reading

and b) reading and producing a written synthesis. The initial hypothesis was that the students who read the texts and produced a synthesis of them would achieve better comprehension outcomes than those who were asked only to read the texts. A secondary aim was to test whether using paper or using computer software had any influence on the comprehension results; we expected that the students who did the task on paper would obtain better comprehension results than those who did it on a computer.

## Method

All 161 participants (121 women and 40 men) were first-year undergraduates in Psychology at a state-run Spanish university who volunteered to take part in the study.

Table 1: Sample distributed by media and task

Media / Task	Paper and pencil	R & A
Reading	39	40
Reading & Synthesis	43	39

The groups who performed the task on a computer used Read & Answer software (Vidal-Abarca et al., 2011) which enabled data to be collected on the participants' reading and writing processes.

Three complementary texts about the current conception of intelligence, taken from thoroughly reliable sources, were adapted for the study. The first was an argumentative piece making a case for theories upholding the diverse character of intelligence as against unitary theories, and contained 973 words. The second, an expository piece, presented Gardner's theory of multiple intelligences (909 words). The third, which was also expository (611 words), presented Sternberg's concept of successful intelligence. Understanding thoroughly these three texts involves comparing, contrasting and integrating information across them. In addition, the following instruments were produced: a prior topic knowledge questionnaire consisting of 21 items ( $\alpha=.50$ ) and a 39-item test to evaluate the comprehension attained after reading the texts ( $\alpha=.60$ ). The comprehension test included questions of different levels of complexity –superficial understanding ( $\alpha=.53$ ), deep understanding and use of knowledge ( $\alpha=.46$ ). On both the prior knowledge questionnaire and the comprehension test, the students had to indicate whether the statements were true or false, and the total number of correct answers was considered.

Before the tasks were administered, the appropriateness and relevance of the source text topic and the type of instruments prepared were assessed by a panel of experts. A pilot study was also carried out to obtain additional information as to their appropriateness.

The data for both tasks (reading/reading & synthesis) and both media (paper/R&A) were obtained in collective sessions lasting approximately 90 minutes. The sessions for the participants using R&A were held in a computer classroom, and those for participants using paper, in an ordinary classroom. The instruction for the synthesis condition was: "You have to write a text including and, above all, integrating, what is most important in the texts you have read on the topic (which is intelligence)".

Enough researchers were present to deal with any possible problems arising during the performance of the tasks. On both tasks and in both media the participants were allowed to spend as much time as they wanted on reading and on producing their written texts. Only when they decided to answer the reading comprehension test the source texts were made inaccessible.

## Results

To test our hypothesis, we conducted three analyses of variance (ANOVA), with media (paper/R&A) and task (reading/reading & synthesis) as independent variables and global comprehension, superficial and deep comprehension questions as the dependent variables. Significant differences were found on the deep comprehension questions from the comprehension test for media condition ( $F_{(3,157)} = 4.62, p <$

.05,  $MSE = 5,53$ ,  $\eta^2 = .029$ ); the students who did the task on paper achieved higher results than those who did it on the computer. No significant differences were found for superficial and global comprehension.

Contrary to our hypothesis, we found no difference between the comprehension outcomes of the students who did the read-only task and those who also produced a synthesis. This rather unexpected finding will be discussed in the light of the qualitative analyses we are carrying out. These analyses focus, on the one hand, on the participants' reading processes (accessible for the subsample that did the task using computer software) and, on the other, on the characteristics of the written text the participants actually produced when asked to synthesize and compare the arguments in the texts. The preliminary results of these analyses seem to indicate that both variables (the reading process in which the participants engage and the characteristics of the text they produce) are crucial to explain the comprehension outcomes. If these results are confirmed, this will be an argument for the need for detailed analyses of the reading and writing processes in which participants engage, as part of any research on tasks similar to those investigated here in order to determine the scope of its results.

## References

- Boscolo, P., Arfé, B. & Quarisa, M. (2007). Improving the quality of students' academic writing: An intervention study. *Studies in Higher Education*, 32(4), 419-438.
- Nelson Spivey, N. (1997). *The constructivist metaphor: Reading, writing, and the making of meaning*. San Diego: Academic Press.
- Mateos, M., & Solé, I. (2009). Synthesising information from various texts: a study of procedures and products at different educational levels. *European Journal of Psychology of Education*, 24 (4), 435-451.
- Murphy, P. K., Long, J. F., Holleran, T. A., & Esterly, E. (2003). Persuasion online or on paper: A new take on an old issue. *Learning and Instruction*, 13, 511-532.
- Pack, P. (2005). Recent trends in comparability studies: Pearson educational measurement. Retrieved 01.06.12. [http://www.pearsonsolutions.com/downloads/research/TrendsCompStudies\\_rr0505.pdf](http://www.pearsonsolutions.com/downloads/research/TrendsCompStudies_rr0505.pdf)
- Segev-Miller, R. (2007). Cognitive Processes in Discourse Synthesis: The Case of Intertextual Processing Strategies. In Rijlaarsdam, G. (series Ed.) and M. Torrance, L. van Waes & D. Galbraith (Volume Eds.). *Writing and Cognition: Research and Applications* (Studies in Writing, Vol. 20, pp. 231-250). Amsterdam: Elsevier.
- Vidal-Abarca, E., Martínez, T., Salmerón, L., Cerdán, R., Gilabert, R., Gil, L., Mañá, A., Lloréns, A., & Ferris, R. (2011). Recording online processes in task-oriented reading with Read&Answer. *Behavior Research Methods Instruments & Computers*.43, 179-192.



# Towards a Suitable Way to Know What Undergraduates Believe about Reading

*Eva Lordán*  
eva.lordan@ub.edu

*Francesc Salvador Beltrán*  
francesc.salvador@ub.edu

*Isabel Solé*  
isoleg@ub.edu  
University of Barcelona  
Pg. Vall d'Hebrón, 171. 08035, Barcelona, Spain

**Abstract.** This research deals with exploring undergraduate students' reading beliefs and highlights some of the strategies that can be used in the construction of a questionnaire. Once our theoretical framework and the definition of the construct of interest took shape, we used three different information sources to select and create items; one of these sources is related to a qualitative study we carried out to explore preliminary ideas about reading in a sample of 32 participants. Afterwards, we analyzed the content validity of a pool of 74 items using a derivation of the Content Validity Index to calculate interjudge agreement; this analysis led to a preliminary proposal for a 60-item questionnaire. We have been carrying out a pilot study in a sample of 60 participants in order to select the final items and pre-analyze the metric properties of the instrument.

**Keywords:** reading beliefs, psychometrics, higher education

The following work exposes a research project in progress aimed at creating a valid, reliable questionnaire to explore the undergraduate students' reading beliefs, taking into account a complex, interactive, epistemic view of reading (Adams & Collins, 1979; Solé, 1992).

Over the last years, various research has been focused on the study of students' beliefs about the nature of knowledge and how it is acquired as well as the instruments they use to acquire it, such as reading or writing, as a relevant factor which may be useful to explain how students represent and face learning situations (e.g.; Mateos et al., 2011; Schommer, 1990; Schraw & Bruning, 1996). This is a relatively new field of study, not exempt of debate and controversy, both with regard to the nature of these beliefs and the ways to study them. We assume the conceptualization of beliefs as implicit models that people create through our personal experience in social learning contexts (Pozo, Scheuer, Mateos & Pérez Echevarría, 2006; Schraw & Bruning, 1996). These models imply restrictions on the way we behave; it is here where the importance to explore them resides.

The reference study on reading belief research, carried out with undergraduate students, was conducted by Schraw & Bruning (1996). The researchers created the *Reading Belief Questionnaire* (RBQ), a Likert scale based on two different implicit models of reading: the transmissional model (reading is a one-way process in which the meaning is transmitted from the author or the text to the reader), and the transactional model (reading is a dynamic process in which the implication of the reader in order to construct a personal meaning of the text is especially relevant).

The results from this study showed that students with high transactional reading beliefs got a more overall, personal, sophisticated representation of the meaning of the text, and included a wider range of personal and critical arguments in a text they wrote as a response to what they had read, than students with low transactional reading beliefs. These findings are consistent with those from later studies using the RBQ –original or adapted–, for expository, narrative and argumentative text understanding. Beyond these results, Dai & Wang (2007) found that transactional readers tend to consider their previous knowledge and experience and to engage more in reading, while transmission readers seem to barely be involved and give up easily when facing understanding difficulties. Finally, a study carried out in our country (Mateos et al., 2011) shows that when students are asked to write a synthesis after reading two argumentative texts, those who hold transactional reading beliefs go beyond

“the given” in the sources, integrating controversial information from the source texts and including critical responses in their own written product.

Together, the aforementioned studies provide arguments to consider that the processes and results related to reading comprehension are sensitive to the beliefs that the reader holds when facing reading, and that the existing instruments for measuring beliefs are relatively appropriate to explore them. However, the internal consistency results from the application of a Spanish adapted version to the RBQ in our country showed that both the transmissional and the transactional beliefs scales were on the limit of being able to be interpreted ( $\alpha=.55$  and  $\alpha=.52$  respectively, Mateos et al., 2011). From our theoretical framework, some conceptual limitations add to the previous ones. The two implicit models in which the structure of the RBQ is based take into account some reading aspects, related to the message and intentions of the author and the implication of the reader in the construction of a personal text meaning. Nevertheless, they disregard other aspects we find necessary to be considered from the complex, interactive reading view that we assume, such as metacognitive factors involved in the assessment and regulation of the reading process and results or the mediated nature of reading from writing. Also, the reading planning and the meaning construction processes are partially covered, especially focusing on the transmissional model and less on the transactional one. These methodological and conceptual limitations offer us arguments to consider the necessity to elaborate a new questionnaire in order to explore reading beliefs.

### **Construction process of the reading beliefs questionnaire**

#### *Focusing on the content validity of the questionnaire*

*Developmental stage.* In order to make the construct of interest measurable, we carried out a thorough review of specialized literature in order to clearly define the construct and specify its dimensions of content operationally. The different perspectives about reading considered in the existing literature (Hoffman, 2009; Pressley & Gaskins, 2006) offer us elements to elaborate our questionnaire based on two reading beliefs viewpoints: a reproductive viewpoint, according to a simple view of reading, and an epistemic viewpoint, according to a constructive view. Several conceptual and/or experimental proposals related to reading and reading comprehension (e.g. Pressley & Gaskins, 2006; Solé, 1992; Van Dijk & Kintsch, 1983) give support to the idea of considering two reading dimensions: the students beliefs about reading as a procedure, on one hand, and about the use of reading to acquire knowledge, on the other.

As operational indicators of the dimensions of content for the construct, we develop items following three different steps. Firstly, we carried out an exhaustive revision on some instruments used in previous studies to measure reading, writing and epistemological beliefs in order to select those items closer to our interest and add them to our questionnaire. Secondly, we created new items from the answers that a group of 32 Spanish undergraduate students –females and males- gave to a 12 open-question format questionnaire about reading, which was created *ad hoc* for this study. We followed an inductive-deductive method to create a category system for each question in order to carry out a content analysis on the answers. Thirdly, we have also created new items from our theoretical framework. After these three steps, an initial pool of 74 items was proposed.

*Judgment stage.* To assess what extent each item was relevant for the construct, the initial set of 74 items was evaluated by a panel of content experts. In order to systematize the evaluation, we generated a set of instruments to assess both the content and the format of the questionnaire. We conducted an interjudge agreement analysis of the experts' answers, using a derivation of the Content Validity Index - CVI (Wynd, Schmidt & Schaefer, 2003), to decide which items should be preserved ( $CVI > .70$ ) or eliminated ( $CVI < .70$ ). The qualitative information provided by the judges was jointly taken into account. The data obtained from the judges' assessment led to reducing the initial pool to 63 items.

### *Focusing on the reliability of the questionnaire responses*

*Pilot study.* We have been carrying out a pilot study of the 63-item version of the questionnaire with a sample of 152 undergraduate students in order to obtain the final version of our reading beliefs questionnaire. Some preliminary results from this study will be provided at the conference.

A strong point of the methodological process we have described is that it can be followed and adapted to construct or adapt reliable, valid instruments to measure a large set of constructs in social sciences and, specifically, in educational research. This process is a compendium of key points that researchers may use as a reference or an aid to face and overcome the difficult process of linking abstract concepts with observable and measurable indicators, in order to make them explicit and tangible to work on and work with.

### **References**

- Adams, M. J., & Collins, A. M. (1979). A schema-theoretic view of reading. In R.O. Freedle (Ed.), *New directions in discourse processing* (pp. 1-22). Norwood, NJ: Ablex.
- Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, 32(3), 332-347.
- Hoffman, J. V. (2009). In search of the “simple view” of reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of Research on Reading Comprehension* (pp. 54-66). New York, USA: Routledge.
- Mateos, M., Cuevas, I., Martín, E., Martín, A., Echeíta G., & Luna, M. (2011). Reading to write an argumentation: the role of epistemological, reading and writing beliefs. *Journal of Research in Reading*, 34(3), 281-297.
- Pressley, M. & Gaskins, I. (2006). Metacognitively competent reading comprehension is constructively responsive reading: How can such reading be developed in students? *Metacognition and Learning*, 1, 99-113.
- Pozo, J. I., Scheuer, N., Mateos, M., & Pérez Echevarría, M. P. (2006). Las teorías implícitas sobre el aprendizaje y la enseñanza. In J. I. Pozo, N. Scheuer, M. P. Pérez Echevarría, M. Mateos, E. Martín & M. de la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 95-132). Barcelona: Graó.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498-504.
- Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading Research Quarterly*, 31(3), 290-305.
- Solé, I. (1992). *Estrategias de lectura*. Barcelona: Graó.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *The Western Journal of Nursing Research*, 25(5), 508-518.

# Using Vibration to Guide Exploration of Haptic Graphics

*Richard K. Lowe*  
Curtin University  
Australia  
r.k.lowe@curtin.edu.au

**Abstract.** Haptic guidance offers a non-visual way of cueing appropriate direction of learner attention in complex animated displays. This methodological paper describes a haptic panel capable of providing vibration in specific locations to guide search or signal target elements. The panel allows the effectiveness of haptic arrays with and without vibration to be compared. A measure of search efficiency in terms of panel area explored was developed using video records of participants' hand movements. Construction and operation of the panel are described together with the approach used to characterize the effects of vibration on tactile search.

**Keywords:** complex animation, visual attention, haptic cueing, vibrational signalling, tactile search

## Introduction

A persistent challenge in using animations to explain complex topics is to direct learner attention reliably to key information within the presentation. The transitory nature of animations means that such visually rich dynamic displays can be very demanding for learners to process (Lowe, 1999). This results in them applying their limited visual attention capacity in a highly selective manner. Unfortunately, learners who are novices with respect to the depicted content tend to direct their attention largely on the basis of its perceptual salience rather than its thematic relevance. Attempts to ameliorate this situation using visual cues have met with mixed success (e.g., de Koning, Tabbers, Rikers, & Pass, 2011). This is perhaps understandable because the addition of further visual signals to an already complex display can make the situation worse rather than better.

An alternative approach could be to use non-visual cueing so that the visual processing channel is not cluttered even further with additional information. One possibility is to accompany the animation with a verbal narration intended to direct learner attention to the high relevance, low salience information that may otherwise be missed. However, verbal representations of spatial and temporal information are relatively imprecise, making provision of well targeted guidance for visual attention in a sufficiently parsimonious format difficult or impossible. The fundamental problem is one of producing a close match between the form of the provided cueing and the visual display within which attention must be directed. Our sense of touch may be able to provide a closer match than can be provided verbally. There are notable parallels between the visual and haptic systems and these offer some potentially useful cross-modal interactions (c.f. Grabowecky, List, Iordanescu, & Suzuki, 2011). However, there are also some key differences, such as sensory resolution and search processes (Lowe & Keehner, 2010; Keehner & Lowe, 2010). This raises fundamental questions about how any proposed haptic guidance system could best deliver its cues and how its fit with the visual system could be optimized. The work reported here involved the design and testing of a haptic tablet for delivering vibrational cues to complement tactile search for test geometric objects based on static properties such as size and shape.

## Haptic tablet design

Under normal conditions, people using their hands to perform tactile search of a given area would need to perform exhaustive scans of that area in order to ensure the search was effective. The time lag produced by such scanning would rule out using conventional tactile graphics for cueing visual animations. However, the application of selective vibration to either the search target or locations leading to the target could make search exploration more efficient by considerably reducing the area needing to be searched. The haptic tablet discussed here has a 4 by 5 grid providing 20 possible active

vibration sites (Figure 1). Each of these sites surmounts a tactor (i.e. electrically controlled vibration cell) whose activation state and vibration intensity are under software control. Vibration from the tactors is transmitted to the surface of the tablet via a square cylinder/piston arrangement. A site can be (i) left empty (ii) fitted with a shaped tile that stands proud of the tablet surface, or (iii) filled with a dummy piston flush with the surface. Sites with a tile or dummy piston can be programmed either to be still or to vibrate.

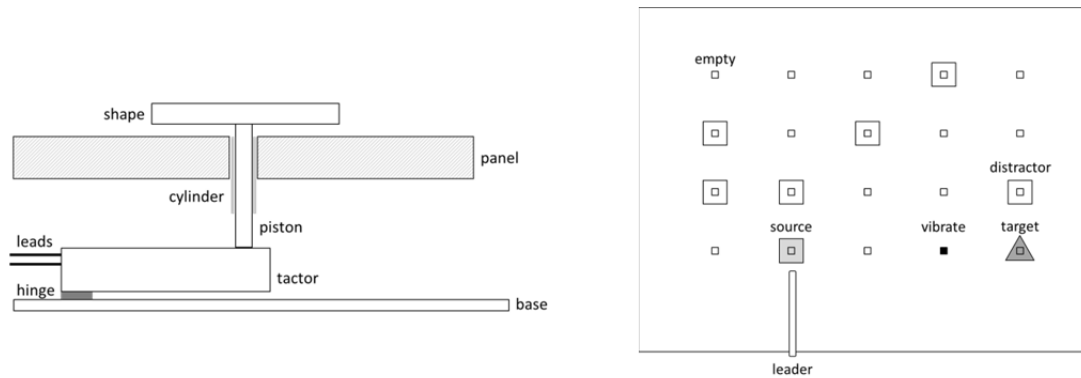


Figure 1. (a) Tablet cross section showing tactor (b) Haptic tablet layout (top view)

### Search under still and vibrating conditions

It was expected that vibration would facilitate search for target shapes on the haptic panel with both area searched and the time to locate the target being smaller than in a still condition. This expectation was tested by presenting blindfolded participants with tasks in which they searched for a target tile that was shaped either the same or differently from a reference tile (the source). The tablet was designed to allow investigation of three haptic exploration conditions: (i) no vibration (ii) vibration applied to the target tile, and (iii) vibration applied between source and target. Previous work indicated that square and triangular tile shapes provided an appropriate level of difficulty for such haptic search. For each task, the participant's hands were first guided to the source tile via a leader rod (see Figure 1b). The participant then searched for the nominated target and indicated when it had been found. A camera located directly above the tablet videoed the participant's hand movements during each search task. To prevent possible aural cues in the vibration conditions, all participants wore sound proofing ear muffs through which masking noise (rain) was delivered. Each participant performed a total of eight search tasks for haptic graphic displays of increasing complexity. After completing the final task, participants' blindfolds were removed and they were asked to reproduce the final (most complex) display they had searched by choosing appropriate cardboard shapes then positioning them on a sheet marked with a 4 by 5 grid.

### Data analysis

The video records captured data about both search times and exploration behaviour. To provide a measure of spatial search efficiency, the video showing hand movements over the course of each task was converted into individual frames. These frames were sampled (e.g. every 5<sup>th</sup> frame) to provide a subset that was manageable to analyse yet accurately reflected the area searched. Image manipulation and analysis software was then used to combine these frames then measure the total area of the tablet over which the participant's hands moved. The final composite of the combined frames was overlaid on the original video to check that it accurately reflected the search coverage. Figure 2 illustrates stages in the analysis process.

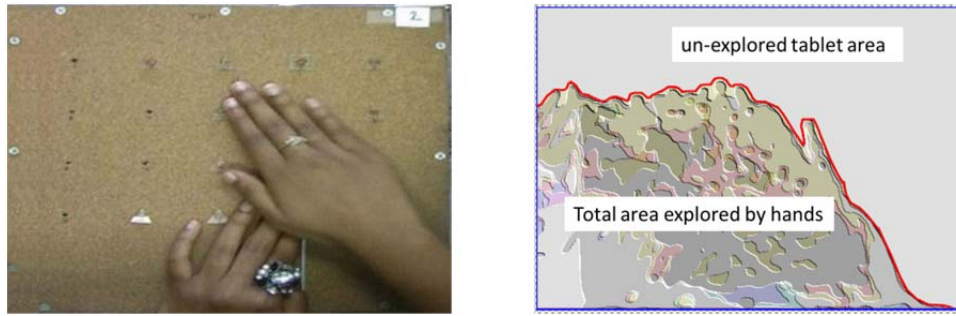


Figure 2. (a) example single frame from search video (b) composite showing total search area

### Interim results and prospects

The most obvious difference between visual and haptic search is that our two hands can carry out independent exploration of a display whereas our eyes operate together. This manual independence was reflected in the search behaviour of participants in both still and vibration conditions. In most cases, search of the tablet area was partitioned between the two hands. Typically, progress of the hands across the tablet was related to the position of the source tile and initially involved generalized sweeps until candidate aspects were found. Exploration then became more focused with the two hands being quite differentiated. For example, one hand might be used as a reference point while the other was used to identify a particular shape. Preliminary indications are that while the use of vibration guidance between source and target dramatically reduced the area searched, it did not have the expected effect of reducing search times. This appears to be because participants in this condition tended to trace very deliberately along the vibration path and which slowed down their location of the target. In the final reproduction task, participants in the vibration conditions tended to generate less complete versions of the tactile graphic displays that typically included only those shapes on or near vibration sites. The indications thus far are that the haptic tablet described here provides a useful tool for comparing how people search still and vibrating tactile graphic displays. It appears that effects of adding vibration cues to guide search may not be uniformly positive, particularly with respect to search duration and recall of the information presented by the display.

### References

- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2011). Attention cueing in an instructional animation: The role of presentation speed. *Computers in Human Behavior*, 27, 41-45
- Grabowecky M, List A, Iordanescu L, Suzuki S, (2011), Haptic influence on visual search, *i-Perception* 2, 824.
- Keehner, M. & Lowe, R. K. (2010). Seeing with the Hands and with the Eyes: The Contributions of Haptic Cues to Anatomical Shape Recognition in Surgery. *AAAI Spring Symposium: Cognitive Shape Processing*.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, 14, 225–244.
- Lowe, R. K., & Keehner, M. (2010, August). Comparing search of tactile and visual graphics. *SIG2 - 2010: Comprehension of Text and Graphics. Tracing the mind: How do we learn from text and graphics?* Tuebingen Germany: Knowledge Media Research Centre.

# Addressing Challenges of Biological Animations

Richard K. Lowe<sup>1</sup>, Jean-Michel Boucheix<sup>2</sup>

<sup>1</sup>Curtin University; <sup>2</sup>University of Burgundy

<sup>1</sup>Perth, Australia; <sup>2</sup>Dijon, France

r.k.lowe@curtin.edu.au; Jean-Michel.Boucheix@u-bourgogne.fr

**Abstract.** Much research on learning with animations has used mechanical rather than biological subject matter. Interactions in mechanical systems tend to involve components acting as coherent wholes (i.e., extrinsic changes such as translations). In contrast, when parts of a biological system interact, they can also undergo transformations (intrinsic changes) due to the non-uniform, non-rigid nature of their constituent material. The behaviour of unfamiliar mechanical systems can therefore be much easier to predict than that of unfamiliar biological systems. This theoretical paper argues that learners may have more difficulty in processing animations of biological systems because of their relative dynamic indeterminacy. It suggests that instead of presenting the animated subject matter in its entirety, key aspects could be dealt with incrementally and cumulatively.

**Keywords:** animation, biological and mechanical systems, transformations, relation sets

## Introduction

To date, subject matter featuring biological systems has been less popular than non-biological topics for investigating learning with animations. With some exceptions (e.g. de Koning et al.), the topics chosen for animation research have tended to involve direct causal chains such as those found in a toilet cistern, or a piano mechanism (Kriz & Hegarty, 2007; Lowe & Boucheix, 2008). These mechanical systems are composed of rigid components whose individual behaviours are regular and predictable from simple physical principles. The functionality of mechanical systems arises from the interaction of these components whose proper operation (via *extrinsic* changes) relies on their form staying essentially constant across space and time. For example, a conventional lever will operate properly only if its shape and size remain fixed. However, the situation with biological systems can be very different. Properties such as flexibility, elasticity and compressibility (as opposed to rigidity) are common for biological materials and these properties often vary across a single component. Non-rigid, non-uniform biological components can undergo all manner of transformations (*intrinsic* changes) in which their shape, size and other characteristics vary in complex ways. Such transformations are typically fundamental to the functioning of biological systems (rather than being an impediment, as they would be in most mechanical systems). However, the associated indeterminacy in response makes the consequences of interactions between biological system components much harder to predict than those of mechanical system component interactions.

For non-rigid entities of the type found in biological systems, behaviour often cannot be readily predicted on the basis of the straightforward dynamics of most objects in our everyday world. For example, although we can be fairly confident that the set of interconnecting levers comprising a piano mechanism will always interact in the same ways, this is not at all the case for biological entities. Figure 1 shows an example of a ‘Sammy’ diagram that depicts the human vocal system. Educators use this type of depiction to help students understand how to pronounce unfamiliar sounds when learning a new language. In order to produce different articulations, the biological entities involved (organs of speech) are transformed in various ways. During speech, multiple transformations of these entities occur simultaneously in different regions. Although Sammy diagrams were originally static graphics, recently they have been converted into animations on the assumption that explicit depiction of speech act dynamics would better support the learning of difficult pronunciations. As is typical with most educational animations, these animated Sammy diagrams provide a comprehensive depiction of the subject matter and its dynamics. However, recent research on learning from animations suggests that presentation of a working system in its entirety can result in students missing some key aspects of the action. This tends to be particularly problematic where multiple events are distributed across the

display so that there is competition for the learner's attention by different dynamic sites (as in the case of an animated Sammy diagram).

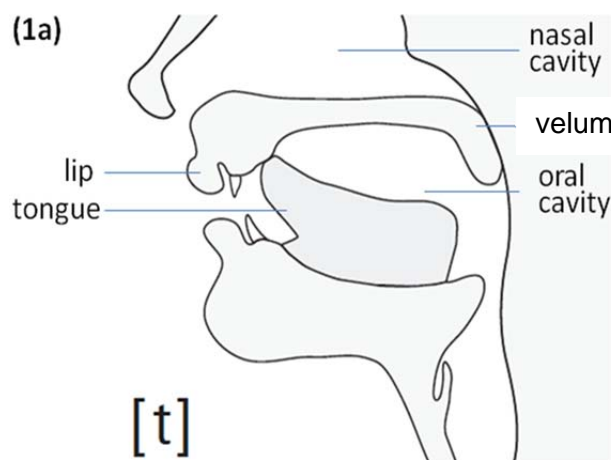


Figure 1. Sammy diagram

### Comparison of processing demands

The relative predictability of component interactions in a mechanical system could facilitate learner processing, particularly with respect to information search. Partial rather than comprehensive learner inspection of a mechanical animation may suffice because of learners' capacity to determine events in locations some distance from where attention is being directed. Determining what happens in currently non-fixated parts of the display could occur by means of top down inference (e.g., applying informal rules about how levers behave) and/or from information collected via peripheral vision. For example, a fixation on one end of a lever that determines its direction and speed of movement allows prediction of how the other (non-fixated) end of the lever will behave. Bottom up information from peripheral vision can confirm or modulate such a prediction. These opportunities make it possible for learners to allocate attention efficiently and strategically because there is no need to examine all parts of the lever in detail. Relieved of the need to search for relations *within* a component, learners can distribute their attention across the display to extract key functional relations *between* components.

However, similar processing 'short-cuts' are less likely to be available for non-rigid components of biological systems. A change in one part of an organ does not necessarily result in a simple, readily predictable change in distal part of that organ. Instead, there are many possibilities as to what type of associated change will take place. Animated Sammy diagrams illustrate this issue. The tongue plays a central role in the production of different sounds by varied changes in its size, shape and position within the mouth cavity. The properties of the tongue allow it to be transformed into myriad configurations, each of which plays a key role in producing particular vocal effects. For language learners to generate a difficult vocalization, they must be able to build a high quality mental model of the necessary tongue transformations. In contrast to the situation with mechanical system animations, extracting all the relevant information from an animated Sammy diagram requires a relatively *comprehensive* inspection of the targeted biological component. Further, in a conventional animated Sammy diagram these complex changes in the tongue are accompanied by numerous visually subtle transformations in the other organs of speech that together produce a particular articulation.

### Conclusion

A need for more thorough analysis of the changes that occur in conventional animations of biological systems coupled with competition for attention makes it probable that learners will miss crucial high



relevance, low salience information. A composition approach using *Relations Sets* (Lowe & Boucheix, in press) has been proposed as a more effective way to supply learners with the raw material they need for building high quality mental models of demanding dynamic subject matter. Relation sets are small groups of functionally-related event units (entities plus their associated behaviours) that are subsets of the target subject matter capable of being assembled incrementally and cumulatively into larger relational structures. Instead of presenting learners with a conventional animation depicting the subject matter in its entirety, a relation set approach presents only one subset of the subject matter at a time. Two key features of a relation set approach to animation design are that it (i) relieves learners of the burden of decomposing an entire animated depiction, and (ii) focuses upon the spatial and temporal relations that are essential for linking key dynamic information into a coherent whole. Connection of event units both within a relation set (*intra-set* links) and between different relation sets (*inter-set* links) are addressed. One approach to developing relation set animations is for the instructional designer to start with a conventional animation that presents the content as a whole and then systematically decompose that material. This decomposition is made not on the basis of the content structure alone (as is done with conventional segmentation) but rather according to relationships that the learner needs to internalize in order to construct a high quality mental model. These relationships are established by analysing the task of internalizing an external dynamic depiction in terms of the Animation Processing Model (Lowe & Boucheix, 2008) which conceptualizes learning from animation in terms of five processing phases. Each of the relation sets devised from this process-based analysis is designed on the assumption that the learner's main activity will be to compose a mental model from subsets of pre-selected raw material chosen to optimize knowledge construction processes. A relation set approach turns the conventional approach to designing educational animations on its head because it specifically targets the internal processes by which learners build mental models (rather than focusing on the external representation of the referent content). Key *within-entity* functional boundaries that non-experts would miss if required to decompose biological content presented as a whole are explicitly identified for learners. This lets them focus their limited processing capacity on the visually subtle transformations that are central to a biological system's functionality.

## References

- de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. G. W. C. (2010b). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20, 111-122.
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human Computer Studies*, 65, 911-930.
- Lowe, R., & Boucheix, J.-M. (2008). Learning from animated diagrams: How are mental models built? In G. Stapleton, J. Howse & J. Lee (Eds.), *Diagrammatic representations and inference – Proceedings of the Fifth International Conference Diagrams 2008* (pp. 266-281). Berlin: Springer Publishers.
- Lowe, R., & Boucheix, J.-M. (in press). Dynamic diagrams: A composition alternative. To appear in P. Rogers, P. Cox & B. Plimmer (Eds.), *Diagrammatic representations and inference – Proceedings of the Sixth International Conference Diagrams 2012*. Berlin: Springer Publishers.

# Visualization and Manipulation of English Sounds on an Interactive Whiteboard at Primary School

*Emilie Magnat*

LIDILEM (Stendhal University, Grenoble III)  
UFR des Sciences du Langage, BP 25, 38040 Grenoble cedex 9, France  
SMART Technologies  
15-17 rue pagès – 92150 Suresnes, France  
emilie.magnat@gmail.com

**Abstract.** Phonemic awareness, the capacity to manipulate sounds independently from the meaning, was examined in the learning of English L2 in French primary school. Such manipulations of sounds are generally done mentally and are demanding on memory, especially in L2. From existing approaches relying on visualization of sounds and gestures, we created activities in which sounds are visualized, can be heard and manipulated. Those activities were done on an Interactive Whiteboard (IWB) with 7-8 year old French-speaking learners. These findings suggest that visualization helps learners getting aware of the importance of phonemes and paying more attention to sound differences while listening to English.

**Keywords:** visualization, phonology, phonemic awareness, interactive whiteboard (IWB), primary school

## Introduction

This research is based on the Kanta and Rey's (2003) study that showed a lack of phonological awareness in English for 11 year-old learners entering secondary school. From a diagnostic based on official instructions, teachers' expectations and a theoretical framework, we designed activities aiming at developing phonological awareness among French speakers working in full class sessions. These activities rely on sound visualization and manipulation to enhance the mental manipulation.

## Diagnostic

According to Bryant and Bradley's (1983) study, phonemic awareness determines future literacy capacities. Lundberg, Olofsson and Wall (1980) went even further proving that it is phoneme manipulation skill that is the strongest predictor of reading and spelling abilities. In their study, Chiappe and Siegel (1999) showed that learners of English L2 are able to transfer their phonological awareness skills in one alphabetic language to develop phonological awareness in English. Other studies support this finding (Comeau, Cormier, Grandmaison and Lacroix, 1999) and show that phonological awareness is not language specific. However, Kanta and Rey (2003) enlightened the fact that French speakers consider English as an expansion of French and so, lack of an operational phonological awareness. In the progression proposed by the National Ministry of Education (Ministère de l'Éducation nationale, 2007), the learning of a second language must be based on an oral approach paying attention to the importance of phonology. However, in the various discussions we had with primary school teachers, it appeared that they don't know how to deal with the section entitled "phonology" in the official instructions. Course books mainly allow teachers and pupils to work on auditory discrimination. But as far as we know, no course book proposes phonological awareness activities that is to say activities of sound manipulation independently from meaning.

## Activities design

Keeping in mind the theoretical framework and teachers' expectation, six sessions of phonemic awareness have been designed. They are integrated into an already existing sequence from a course book that is currently used in class.

### *Modernization of existing methods*

In the 1970s, educators like Gattegno, Borel Maisonnay Sue Lloyd and Sara Wernham designed learning approaches relying on visual aids and gestures. The "Look and say" approach was popular at

the time. Dr. Caleb Gattegno (1991-1988) developed the *Silent Way* for the learning of foreign languages. Instead of using phonetic spelling, he used colour-coded instructional materials. In this approach, learners can use the *Fidel* (“phonic code”) and coloured words to progressively match the sound system of a language with its written system (Gattegno, 1978). Borel Maisonnay’s method is closed to Lloyd and Wemham’s *Jolly Phonics approach*. Both methods rely on gestures and were originally designed for special need learners. The *Silent Way* as well as *Jolly Phonics* and Borel Maisonnay’s approach rely on visual aid to enhance the discovery of the sound structure of a language.

### *The activities created*

So that teachers and learners might visualize and physically manipulate sounds, we created coloured cards based on the same palette of colour used by Gattegno. Each card gets its respective recorded sound joined so that when touching the card everybody can hear the corresponding sound. These cards can be moved on the board, still keeping the sound linked. Hence, these cards are external representations of these mental items that are the sounds. Cards can be manipulated and heard as often as needed and so being a cognitive help to realize phonological awareness activities.



Figure 1. Colored-cards used without phonetic signs



Figure 2. Example of activities where cards should be manipulated

## **Evaluation of the activities**

### *Protocol of research*

To determine the impact of phonological activities done on an Interactive whiteboard using visualization, two groups were compared. The first group used the created activities on the IWB while the second group had the same activities to do with the same items but without the interactive whiteboard. The latter is part of a system composed of a videoprojector and a computer. It is a board on which information coming from the computer is projected, but it is also a input/output device from which teachers and pupils can take control of the computer. The second group used a projected powerpoint presentation on which representations of the items were presented but no visualization or manipulation of phonemes was proposed. This group also had the recorded sound, but on a CD that the teacher could use on a CD player. Hence, the second group carried out activities of phonological awareness on a traditional manner that is to say only mentally while the first group had the IWB as an help providing visualization and allowing manipulation of phonemes. During the six sessions, both groups were observed. We kept a record of the various actions that were done on the IWB as well as the different discussion-time between learners and the answers they provided. Both groups were tested before and after the activities on their phonological awareness, auditory discrimination and phonological memory in French as well as in English. Furthermore, both groups and teachers were interviewed afterwards. To get a precise analysis, we triangulated information coming from the teacher, pupils and the observer.

### *Quantitative results*

A linear regression controlling the pre-test results and taking into account the interactions between groups shows that the group using the IWB had a better progression than the without-IWB group on the total result of the phoneme reversal test [ $R^2=0.776$  ( $F=61.058$ ,  $p<0.0001$ )]. But on the other hand, the without-IWB group got a better progression as far as phonological memory is concerned [ $R^2=0.336$  ( $F=5.172$ ,  $p<0.001$ )].

### *Qualitative results*

*On pupils' perspective.* While using those activities, learners using the IWB realized how important it is to pay attention to sound distinctions. Within a remarkable short amount of time, learners were able to say to which sound a particular coloured card refers. Learners who did not use the IWB activities had a struggle to make a chain of three English sound words while the IWB class was more aware of the characteristics of English sounds. Besides, while observing the two classes we also noticed a difference of attention. In the IWB class, every pupil was looking at the board trying to do the task even if it wasn't their turn. This attention was not that high in the class without the IWB where materials were static all the session long. We can hypothesize that it is the manipulations that were done on the board that catch pupils' attention. Moreover, during the pretests and posttests, learners were asked to mentally manipulate sounds. For instance, in the phoneme reversal task they had to say all the sounds of a word in the other order. During the posttest, we noticed that some learners from the IWB class made some movements on the table or counted on their hands. We tried to make them verbalize this during the interview but none of them were able to explain such an attitude. However, the teacher thinks that it might be due to the manipulating activities done on the IWB.

*On the teacher's perspective.* During her interview, the teacher using the IWB admitted that, although at first she found the activities difficult for her pupils, she was really proud to see that they coped on and succeeded. She also felt more at ease to deal with the phonological aspect in learning a second language with a tool like an interactive whiteboard. She used the latter as a cognitive help to pronounce words: she listened to the board and then could repeat the word paying attention to stick to the example she got and so, she provided the right input to learners.

### **Conclusion and perspectives**

In a nutshell, the interactive whiteboard allows us to create activities in which sounds can be visualized, heard and manipulated. These external representations seem to have been a successful help for the teacher. It also seems to have helped learners developing their awareness of the sound structure of English. Those qualitative results will be analysed in more details along with the quantitative results in order to get a more precise view of the impact of such activities. At the wake of the results, we might be able to provide teaching guidelines and improve teaching materials.

### **References**

- Bryant, P., & Bradley, L. (1983). Categorizing sounds and learning to read: a causal connection. *Nature*, 301, 419–421.
- Chiappe, P., & Siegel, L.S. (1999). Phonological awareness and reading acquisition in English- and Punjabi-speaking Canadian children. *Journal of Educational Psychology*, 91(1), 20–28.
- Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 91(1), 29–43.
- Gattegno, C. (1978). *The Silent Way : teaching foreign languages in schools*. New York: Educational Solutions Worldwide Inc.
- Kanta, T., & Rey, V. (2003). Relation entre la conscience phonologique et l'apprentissage d'une langue seconde. *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 22, 135–147.
- Lundberg, L., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology*, 21, 159–173.
- Ministère de l'Éducation nationale. (2007). Programme de langues étrangères pour l'école primaire anglaise. *B.O. Hors Série du 30 août 2007*, 20–33.

# What Features Make Decorative Illustrations Interesting?

Ulrike Magner

Rolf Schwonke

Inga Glogger

Alexander Renkl

Educational and Developmental Psychology, University of Freiburg, Germany  
ulrike.magner@psychologie.uni-freiburg.de, rolf.schwonke@psychologie.uni-freiburg.de,  
inga.glogger@psychologie.uni-freiburg.de, renkl@psychologie.uni-freiburg.de

**Abstract.** In order to study the influence of situational interest on learning, an instructional designer might want to know what features make decorative illustrations interesting. To answer this question we analyzed situational interest triggered by decorative illustrations and the relevance of 4 features: concreteness, personal relevance, unexpected information, and ease of comprehension. We used multilevel analysis to fully control for inter-individual differences in “third variables” such as individual (“habitual”) interest. Our results showed that concreteness, personal relevance, and ease of comprehension are essential features of decorative illustrations to trigger situational interest. Unexpected information did not make unique contributions in explaining how interesting different illustrations are perceived. The present study provides foundational knowledge for the “motivational design” of learning environments.

**Keywords:** Interest research, illustrations, mathematic learning, multilevel analysis

Imagine your participants rate several *decorative* (defined as showed little or no relationship to the text content; Carney & Levin, 2002) illustrations, each embedded in a geometry task, with respect to *situational interest* as well as to four potentially interest-triggering features, that is, concreteness, ease of comprehension, personal relevance, and unexpected information. How can you know whether interrelations between these features and *situational interest* triggered by decorative illustrations are “substantial” or just due to *individual* (“habitual”) *interest* or other third variables such a prior knowledge? The present study contributes to answering this question.

*Situational interest* is defined as a reaction to environmental input, for example, decorative (decorative) illustrations (Hidi, 2001). In contrast, *individual interest* is a relatively stable predisposition which has emerged over time (Hidi, 2001). Situational and individual interests are related. Research on situational interest in texts has shown that in particular concreteness, ease of comprehension, personal relevance, and unexpected information trigger situational interest (Schank, 1979; Sadoski, Goetz, & Rodriguez, 2000).

However, we think that the empirical evidence on the relations between situational interest and interest-triggering features is sub-optimal. First, most of the research on situational interest refers to text learning (e.g., Sadoski et al., 2000) and not do learning with illustrations. Second, an experimental approach (e.g., Harp & Mayer, 1997; Sadoski et al., 2000) that varies one of the potentially interest-triggering features such as personal relevance has the disadvantages that it is hardly possible to find a feature that is personally (equally) relevant for all learners; in addition, it is hard to vary such features as personal relevance without changing other characteristics as well (e.g., type of content). Third, when correlational techniques are applied (e.g., path analysis; Sadoski et al., 2000), it is highly plausible that relations between situational interest and features such as personal relevance are in part due to differences in individual interest. That means that individual interest (or even other variables such as prior knowledge) determines both the situational interest and the rating of interest-triggering features such as personal relevance; hence the correlations between situational interest and personal relevance might be inflated.

In order to overcome these problems, we analysed the relations between the four potentially interest-triggering features of illustrations (concreteness, ease of comprehension, personal relevance, and unexpected information) and situational interest on an intra-individual basis. Thereby, the

influence of potentially confounding variables such as individual interest or prior knowledge is excluded. For this purpose we used multilevel analysis.

## Method

We used the data of a previous study (Magner, Schwonke, Renkl, Aleven, & Popescu, 2010). Participants were 87 eighth grade students (35 male; 52 female; age:  $M = 13.9$  years;  $SD = 0.6$ ) from a German secondary school. They rated their situational interest triggered by 11 (out of a pool of 44 illustrations) randomly assigned decorative illustrations on nine-point scales. Therefore we had 957 ratings of situational interest triggered by decorative illustrations. The decorative illustrations were embedded in screenshots from a multimedia learning environment on intersecting lines (geometry). The illustrations showed real-life situations or objects such as sitting in front of a computer, cycling, or a compass. They were not directly related to the content of intersecting line problem. Furthermore, the students rated each decorative illustration's concreteness, ease of comprehension, personal relevance, and unexpected information on nine-point Likert scales.

In order to analyse the relations between potentially interest-triggering features of decorative illustrations and situational interest on an intra-individual basis controlling for confounding variables such as individual interest or prior knowledge, we used multilevel analysis (hierarchical linear modelling; HLM; Raudenbush, Bryk, & Congdon, 2008). Multilevel analysis is useful because students rated several decorative illustrations, that is, our data (957 illustration-ratings) are nested in individuals. Typically, hierarchical linear modelling is applied in situations in which students within classrooms are more similar, for example, in their learning outcomes than students in different classrooms. Multilevel analysis overcomes this dependency of data points within classrooms by modelling classrooms on a second level (and individuals on level 1). Although done less frequently, multilevel analysis can also be used for modelling data within individuals and simultaneously taking between-person variations into account. We used multilevel analysis in this way. More specifically, situational interest was predicted by including the ratings of concreteness, ease of comprehension, personal relevance, and unexpected information on the *intra*-individual level 1 ( $n = 957$  illustration-ratings). *Inter*-individual differences in ratings were controlled by modelling individuals on a second level ( $n = 87$ ).

## Results

The first step in our multilevel analysis was to determine the baseline estimation of the variance within and between individuals (Raudenbush & Bryk, 2002). For this purpose, we only included the dependent variable situational interest triggered by decorative illustrations in an HLM model (i.e., the null-model). Nearly 55% of the total variance of situational interest was explained by inter-individual differences in rating of decorative illustrations. This result shows that it is sensible to conduct multilevel analysis in order to control for the very substantial individual differences. Controlling for these individual differences in situational interest, we were interested in the relationship of the illustrations' concreteness, ease of comprehension, personal relevance, and unexpected information with situational interest. All potentially interest-triggering features significantly predicted intra-individual variation in how interesting different decorative illustrations were rated (HLM coefficient: concreteness .12,  $t = 3.44$ ,  $p = .001$ ; personal relevance .24,  $t = 6.70$ ,  $p < .001$ ; ease of comprehension .11,  $t = 2.85$ ,  $p = .005$ ), except of unexpected information (HLM coefficient: .02  $t = 0.36$ ,  $p = .720$ ). Concreteness, ease of comprehension, and personal relevance explained 24 % of the error variance on level 1 (i.e., intra-individual level).

## Discussion

Methodologically, we found that situational interest in geometry varied substantially as a function of individual differences (55% of the variance was explained on the inter-individual differences). Therefore, it was sensible – and might be for future studies – to use multilevel analysis for determining the relations between situational interest and potentially interest-triggering features.

With respect to triggering situational interest we found that such interest is elicited by decorative illustrations that are concrete, have personal relevance and ease comprehension. Note that all these features uniquely contributed to situational interest. Hence, instructional designers should use all corresponding “set screws” when trying to optimize their learning environments in motivational respects.

Unexpected information was – in contrast to findings on texts and interest – not relevant for triggering situational interest. Maybe, in texts unexpected information such as interesting but irrelevant details can more easily be included by anecdotes or jokes, as compared to decorative illustrations in multimedia environments. Maybe, illustrations showed too low unexpected information in the present case and therefore no relationship to situational interest could be found.

Overall, previous findings on potentially interest-eliciting features of texts could be largely replicated for decorative illustrations in multimedia environments. Results of the present study can be used to optimize the “motivational design” of learning environments in research and practice. Further studies should analyse in more detail the role of unexpected information in such environments.

## References

- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14(1), 5–26. doi:10.1023/A:1013176309260
- Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, 13(3), 191–209.
- Magner, U., Schwonke, R., Renkl, A., Aleven, V., & Popescu, O. (2010). Pictorial illustrations in intelligent tutoring systems: Do they distract or elicit interest and engagement? In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010) - Volume 1, Full Papers*. International Society of the Learning Sciences: Chicago IL.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2. ed.). Thousand Oaks, CA: Sage Publications.
- Raudenbush, S., Bryk, A., & Congdon, R. (2008). Hierarchical linear and nonlinear modeling (Version 6.06): Lincolnwood: Scientific Software International.
- Sadoski, M., Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92(1), 85–95.
- Schank, R. (1979). Interestingness: Controlling inferences. *Artificial Intelligence*, 12(3), 273–297.

# Effects of Text-Belief Consistency and Reading Goals on the Comprehension of Multiple Science-Related Texts

Johanna Maier, Tobias Richter

University of Kassel, Department of Psychology  
Holländische Straße 36-38, 34127 Kassel, Germany  
johanna.maier@uni-kassel.de, tobias.richter@uni-kassel.de

**Abstract.** When learners inform themselves about currently debated scientific issue, they will read multiple texts that are more or less belief-consistent. Following the congeniality hypothesis, we predict that prior beliefs bias learners' situation model in favor of belief-consistent texts (belief-consistency effect). Moreover, we assume that learners' reading goal moderates the belief-consistency effect. Prior beliefs are especially relevant if learners follow the goal to build an own point of view (epistemic reading goal) in contrast to memorize as many facts as possible (receptive reading goal). In the present experiment, these assumptions concerning the influence of prior beliefs and reading goals on the situation model in comprehending multiple conflicting texts were investigated. Students read one belief-consistent and -inconsistent text each about a scientific issue while thinking aloud and afterwards, comprehension was assessed. In line with our assumptions, results revealed that learners experienced a belief-consistency effect when they followed an epistemic reading goal.

**Keywords:** multiple texts, text-belief consistency, reading goal

Learners more often than not read several texts on the World Wide Web when they want to inform themselves about currently debated scientific issues. Due to the fast publication of new scientific results and divergent theoretical explanations for scientific incidences, the World Wide Web offers a multitude of documents which often represent conflicting argumentative positions on the same scientific issue. This raises the question how learners are able to construct a coherent mental model of the scientific issue on the basis of conflicting information (learning with multiple texts; Perfetti, Rouet, & Britt, 1999). Moreover, learners will not come across new information with a complete open mind but rather come equipped with prior beliefs that might be closer to one argumentative position in the controversy over another one. The present experiment investigated the effects of learners' prior beliefs and reading goals on the mental representation (situation model) that learners construct in the comprehension of multiple scientific texts. Besides investigating comprehension outcomes, we also collected think-aloud protocols to analyse in depth how learners' reading goals influence the processing of multiple belief-consistent and -inconsistent scientific texts<sup>5</sup>.

## Influence of Prior Beliefs on Multiple Text Comprehension

Learners' prior beliefs can be seen as a knowledge structure or schema that influences the comprehension and retrieval of conflicting arguments. More precisely, research provides evidence that belief-consistent information is integrated more strongly into the mental representation of controversial issues (congeniality hypothesis, Eagly & Chaiken, 1993). This effect is likely to be associated with a higher amount of reconstructive processing such as integrating text information with prior knowledge and beliefs by constructing inferences as it is involved in the construction of a situation model. Taken together we assume that prior beliefs on a scientific issue should serve as a knowledge-based structure similar to a schema and thus, should guide the construction of a situation model of the scientific issue in favour of belief-consistent information (*belief-consistency effect*).

## Influence of Learners' Reading Goal on Multiple Text Comprehension

Not only prior beliefs, but also learners' reading goals influence the processing of multiple scientific texts. In a recent review, McCrudden and Schraw (2007) propose that reading goals impact the mental

---

<sup>5</sup> Think-aloud protocols have not been fully analyzed yet, but will be analyzed by the time of the conference.



representation learners construct of texts by guiding attention to text segments relevant for learners' reading goals. For instance, instructing learners to read for a specific purpose (e.g. study vs. entertainment) influences the amount of inferential activities learners engage in during reading. In learning with multiple texts, the goal of building an own point of view might assist learners in building a balanced situation model of the scientific issue by stimulating inferential activities which connect text information not only within but also across texts. In line, Wiley and Voss (1999) have shown that participants who were instructed to write an argument performed better in inference and analogy tasks after reading multiple texts compared to participants that were instructed to write a narrative, summary or explanation. However, in learning with belief-consistent and belief-inconsistent multiple texts, reading goals that instruct learners to build an own point of view about a scientific issue might not only support learners' inferential activities, but might also prompt learners to evaluate text information in light of their prior beliefs. Such an *epistemic reading goal* can be seen as a special case of reading for perspective (McCrudden & Schraw, 2007). In contrast, a *receptive reading goal* that involves the memorization of text information should lead to less inferential and evaluative activities across texts because it directs learners' focus on individual facts within texts.

### **The present study**

Our main goals of the present experiment were to investigate the effects of learners' prior beliefs and learners' reading goals on the comprehension of belief-consistent and belief-inconsistent multiple texts. In line with the congeniality hypothesis we expected a belief-consistency effect in the situation model of the scientific issue. Moreover, we assumed that an epistemic reading goal should increase the belief-consistency effect, whereas this effect should be weaker under a receptive reading goal.

### **Method**

Twenty-seven university students read two texts about possible health risks caused by electromagnetic radiation of cell phones. The belief-consistent text was in line with prior beliefs of the majority of the German population and the participants in this study and argued against possible health risks caused by electromagnetic radiation. The belief-inconsistent text took the opposite stance and argued in favour of possible health risks. Half of the participants were instructed to build an own point of view about the issue (epistemic reading goal), half were instructed to memorize as many facts as possible (receptive reading goal). Participants were requested to think-aloud during reading. After participants finished reading, situation model strength was assessed with a verification task, in which participants decided for three types of test items (paraphrases, inferences, and distractors) whether or not they represented information that matches the texts content. Situation model estimates were based on the (bias-corrected) proportion of correctly recognized inferences in this task.

### **Results and Discussion**

The hypotheses concerning effects of text-belief consistency and reading goal were tested with an ANCOVA for designs with between- and within-subjects factors. The *text order* (pro-contra vs. contra-pro, varied between-subjects), the *text-belief-consistency* (belief-consistent vs. belief-inconsistent, varied within-subjects) and *participants' reading goal* (epistemic vs. receptive, varied between-subjects) were included in the analysis. To control for effects of reading time, the difference in reading times of the two texts were included as covariate in the analysis.

In line with the congeniality hypothesis, results revealed a text-belief consistency effect for the situation model,  $F(1,22) = 12.4, p < .01, \eta_p^2 = .36$ . However, this main effect was qualified by a three-way interaction of text-belief consistency, reading goal and text order,  $F(1,22) = 11.3, p < .01, \eta_p^2 = .34$  (Figure 1). For learners who followed an epistemic reading goal, the situation model for the belief-consistent text excelled the situation model for the belief-inconsistent text independent of text order.

One interpretation of this finding is that learners deliberately disconfirmed belief-inconsistent information when they were instructed to build an own point of view about the scientific issue (Edwards & Smith, 1996). A slightly different possibility is that learners actively and routinely validate incoming text information against their prior beliefs (epistemic monitoring) and tend to reject information which they found implausible (Richter, 2011).

In contrast, learners following a receptive reading goal constructed a stronger situation model for the text they had read at last, independent of text-belief consistency. This effect might be interpreted as result of the memorization process in which information from the last read text overwrote information from the first text (similar to recency effects). In order to clarify the mechanisms underlying the text-belief consistency effect and its interaction with learners' reading goal, we are investigating think-aloud protocols that have been collected during reading.

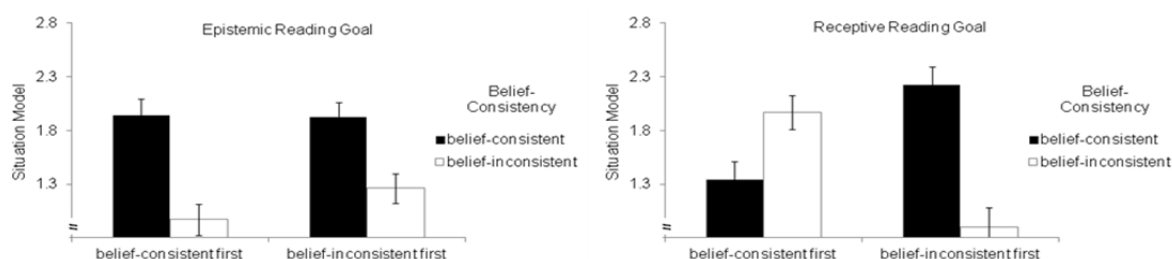


Figure 1. Three-way interaction of text-belief consistency, reading goal and text order for the situation model (Error bars represent the standard error of the mean)

## References

- Eagly, A.H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt, Brace, Jovanovich.
- Edwards, K., & Smith, E.E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71, 5-24.
- McCrudden, M., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, 19, 113-139.
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S.R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99-122). Mahwah, NJ: Erlbaum.
- Richter, T. (2011). Cognitive flexibility and epistemic validation in learning from multiple texts. In J. Elen, E. Stahl, R. Bromme, & G. Clarebout (Eds.), *Links between beliefs and cognitive flexibility* (pp. 125-140). Berlin: Springer.
- Wiley, J., & Voss, J.F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91, 1-11.

# The Role of Epistemological Beliefs and Reading Beliefs in Multiple Text Comprehension

*Mar Mateos*

Universidad Autónoma de Madrid  
Facultad de Psicología, Campus de Cantoblanco, 28049 Madrid, Spain  
mar.mateos@uam.es

*Isabel Solé*

Universitat de Barcelona  
Facultat de Psicologia, Pg. de la Vall d'Hebron, 171, 08035 Barcelona, Spain  
isoleg@ub.edu

*Nuria Castells*

Universitat de Barcelona, Spain  
nuria.castells@ub.edu

*Jara González Lamas*

Universidad Autónoma de Madrid, Spain  
jara.gonzalez@uam.es

**Abstract.** The aim of this study was to examine the relations between epistemological beliefs and reading beliefs, and their role in comprehension when the task consists in integrating information from multiple texts and presenting different positions regarding a controversial topic. A total of 911 university students on different courses and in different years participated in the study. The assessment instruments used were the *EQEBI* (Ordóñez, Ponsoda, Abad & Romero, 2009), to evaluate epistemological beliefs, and the *Reader Belief Questionnaire* (Schraw & Bruning, 1996). Following an assessment of their prior knowledge, the students read three texts about nuclear energy and their intratextual and intertextual comprehension was assessed. The results showed that epistemological and reading beliefs play a moderate part in the comprehension of argumentative texts presenting opposing positions. In addition, the role of epistemological beliefs, which are more general, seems to be mediated by that of reading beliefs, which are more specific.

**Keywords:** epistemological beliefs, reading beliefs, multiple text comprehension, argumentative texts

## Introduction

Reading from multiple texts requires exploring and integrating various sides of an issue and, therefore, may require a particular way of conceiving the nature of knowledge and the nature of reading; in other words, it may be affected by the type of epistemological beliefs and reading beliefs held by the student.

Recent research has looked at the part played by epistemological beliefs in multiple text comprehension (Bräten & Strömso, 2010, Strömso, Bräten & Samuelstuen, 2008) and in understanding texts putting forward different positions on a particular issue (Kardash & Howell, 2000). Epistemological beliefs have been approached from various relatively independent standpoints. However in this study we have specifically adopted the model put forward by Schommer (1990), one of the most influential researchers in this field. To a lesser extent, reading beliefs and their relation to both narrative and expository single text comprehension have also been investigated (Dai & Wang, 2007; Schraw, 2000). Although the two constructs have been explored separately by different traditions, they present certain similarities. Conceptually, they both refer to ways of conceiving of knowledge and meaning, and how knowledge and meaning are acquired, and they are both associated with the degree of personal engagement and elaboration in text comprehension. Both constructs lead to the prediction that students with more complex beliefs will attain higher levels of comprehension than students with simpler beliefs. Based on the results of a previous study (Mateos et al., 2011), this study started out from the assumption that these two kinds of beliefs are not independent, but related. Nevertheless, epistemological beliefs may be regarded as more general than reading beliefs on the understanding that epistemological beliefs would predict text comprehension through reading beliefs.

The aim of the study was to analyse the relations between epistemological beliefs and reading beliefs, and their role in comprehension when the task is to integrate information from multiple texts presenting different positions on a controversial topic.

## Method

The participants were 911 students on different courses (history, philology, engineering, and psychology) and in different years. Epistemological beliefs were assessed using the *EQEBI* questionnaire (Ordoñez, Ponsoda, Abad & Romero, 2009). This new instrument integrates and expands the Spanish versions of Epistemological Questionnaire (EQ) (Schommer, 1990) and the Epistemic Beliefs Inventory (EBI) (Bendixen, Schraw & Dunkle, 1998). EQEBI retains four dimensions: simple knowledge (SK), certain knowledge (CK), quick learning (QL) and innate ability (IA). The students' beliefs about reading were assessed using the *Reader Belief Questionnaire* developed by Schraw and Bruning (1996). They have distinguished the transmission model (TM), involving the belief that meaning must be transmitted from the author and/or the text into the reader's memory and the transaction model (TA), entailing the belief that meaning must be actively constructed by readers integrating their own thinking into the process. Three texts were written on the issue of nuclear energy: an expository text and two argumentative texts presenting different positions. A 20-item test was produced to evaluate prior topic knowledge and a 22-item test was produced to assess intratextual and intertextual comprehension. The latter test was drawn up in such a way that some items could be answered on the basis of information from one of the texts, whereas other items required the students to integrate information from at least two texts. Some items were paraphrases of ideas from the texts, whereas others required the students to make inferences from the information provided by the texts.

## Results and conclusions

A correlation analysis was carried out in order to examine the extent of the relations among all variables. Significant and positive correlations were found between SK and CK with TM and negative with TA. Moreover, SK, CK, and TM correlated negatively with reading comprehension whereas TA correlated positively with reading comprehension.

In order to test the hypothesis of a mediated relation among epistemological beliefs, reading beliefs and reading comprehension, hierarchical multiple regression analyses were performed following the procedure established by Baron and Kenny (1986). The model with the greatest explanatory power (step 5),  $R^2 = .052$ ,  $F_{(2, 845)} = 9.16$ ,  $p = .001$ , showed that, once prior knowledge had been controlled for, transactional reading beliefs contributed to explain variance in comprehension, while the initial contribution of certain knowledge and simple knowledge epistemological beliefs in the previous model (step 4) ceased to be significant. This was due to the fact that the relation between simple knowledge belief and reading comprehension is completely mediated by transactional reading beliefs, as corroborated by the results of the Sobel test,  $z = -1.99$ ,  $p = .023$ .

The findings of this study therefore support the following conclusions, in agreement with the initial hypothesis: on the one hand, epistemological beliefs and reading beliefs are not independent of each other; on the other, epistemological beliefs, which are more general, are related to comprehension through reading beliefs, which are more specific. The results also indicate that beliefs about knowledge and reading play a moderate role in the comprehension of argumentative texts presenting opposing positions and so help to draw the complex map of variables involved in comprehension. In future research, it would be interesting to focus on the role of more specific beliefs, like domain-specific or topic epistemological beliefs as Bråten and others have done in different studies.

## References

- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-82.
- Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic beliefs and moral reasoning. *Journal of Psychology*, 132, 187-200.
- Bräten, I., & Strömso, H.I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, 47, 1-31.
- Dai, D.Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, 32, 332-347.
- Kardash, C.A., & Howell, K.L. (2000). Effects of epistemic beliefs and topic-specific beliefs on undergraduates' cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology*, 92, 524-535.
- Mateos, M., Cuevas, I., Martín, E., Martín, A., Echeíta G., & Luna, M. (2011). Reading to write an argumentation: the role of epistemological, reading and writing beliefs. *Journal of Research in Reading*, 34, 281-297.
- Ordóñez, X.G., Ponsoda, V., Abad, F.J., & Romero, S.J. (2009). Measurement of Epistemological Beliefs: Psychometric Properties of the EQEBI Test Scores. *Educational and Psychological Measurement*, 69, 287-302.
- Schraw, G. (2000). Reader beliefs and meaning construction in narrative text. *Journal of Educational Psychology*, 92, 96-106.
- Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading Research Quarterly*, 31, 290-305.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504.
- Strömso, H.I., Bräten, I., & Samuelstuen, M.S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction*, 18, 513-527.

# Effects of a Mind Map Intervention on Fifth and Sixth Graders' Learning from Texts

*Emmelien Merchie, Hilde Van Keer*

Ghent University

H. Dunantlaan 2, Belgium

Emmelien.Merchie@UGent.be, Hilde.VanKeer@Ugent.be

**Abstract.** From the age of 11-13, children start to spend increasingly more time on learning from texts. The need arises to support them in dealing with text information and engaging them in self-regulated learning (SRL). This study is embedded within the cognitive component of SRL and focuses specifically on Mind Mapping (MM). Several factors play a role, however, in the effectiveness of MM, such as working with author-provided or student-generated MM, cognitive load, and students' general achievement level. Two research questions are addressed: What is the impact of an author-provided or student-generated MM intervention on students' (1) cognitive load and (2) learning strategies, text acquisition, and schematisation skills? Results reveal interesting differences between conditions and students' achievement level.

**Keywords:** Self-regulated learning, Informative texts, Graphic organizers, Mind Mapping, Primary school

## Theoretical Background

Currently, students are continuously challenged by the exponential knowledge increase in our information society and the large amounts of information they have to process and learn independently. This becomes increasingly important from the age of 11-13, when the focus shifts from 'learning how to read' to 'reading to learn' (Bakken & Whedon, 2002). Within this context, the concept of self-regulated learning (SRL) has recently played an important role in educational research (Black, et al., 2006). This study is embedded in the cognitive component of SRL, focusing on learning strategies and tactics students can apply to effectively process and acquire new information (Pintrich, 2004). In this respect, working with Mind Maps (MM) (Buzan, 1974) seems promising in promoting efficient structuring, processing, and learning texts (Farrand, et al., 2002). MM are characterised by specific features (e.g. dimensionality, colours), effective in processing and acquiring information. Also the Dual Coding Theory and Cognitive Load Theory point to advantages in the decrease of cognitive load and the recall of text information when working with maps.

Several factors play a role in the effectiveness of MM to stimulate text processing and learning strategies. First, cognitive load seems to depend on whether MM are author-provided or student-generated (Stull & Mayer, 2007). Furthermore, learner characteristics (e.g. grade and achievement), play a role as well (Vekiri, 2002). Prior research did not always take this into account, leading to a lack of studies focusing on the influence of these characteristics on the effectiveness of MM interventions. This is especially true for primary education, since research on graphic organisers generally focuses on secondary and higher education, although the importance of processing skills in earlier grades is frequently stressed (Rawson & Dunlosky, 2007)

## Research Questions

Two research questions are addressed. What is the impact of an author-provided or student-generated MM intervention on students' (1) cognitive load during the intervention and (2) learning strategies, text acquisition and schematisation skills?

## Research Method

### *Design*

A quasi-experimental repeated measures design was applied in authentic classes. The focus lies on the time interval (pretest and posttest) and the research condition (two experimental conditions: working

with author-provided or student-generated MM and one control condition). Both MM interventions consist of a twelve weeks during MM training (Merchie, 2009). The main goal of both experimental conditions was to stimulate children in processing, acquiring, and learning informative texts in a structured way. In the first experimental condition, these skills were practiced while working with author-provided Mind Maps. In the second experimental condition, students were taught gradually to construct Mind Maps from the texts themselves. The control condition was not explicitly trained in text processing and schematisation skills.

### *Participants*

20 fifth and 22 sixth-grade classes participated ( $n=644$ ), with 213 students in the author-provided MM condition, 219 in the student-generated MM condition, and 212 students in the control condition. Students were divided into high ( $n=291$ ), average ( $n=230$ ), and low achievers ( $n=120$ ).

### *Instruments*

*Perceived cognitive load* was measured in the experimental conditions by the Cognitive Load rating scale (Paas, 1992) and administered after each MM lesson. Pupils were asked to estimate individually how much effort it took them to complete the exercises on a 7-point scale.

Students' learning strategies, text acquisition, and schematisation skills were assessed by a specific learning task administered in class. Students were asked to study an informative text (500 words) for 30 minutes and were free to use a scratch paper. Students were then given a task-specific self-report questionnaire aimed at assessing the use of different *learning strategies*. The questionnaire consists of 41 items, on a 5-point Likert-scale, subdivided into the following subscales: summarising ( $\alpha=.89$ ), highlighting ( $\alpha=.68$ ), rereading ( $\alpha=.74$ ), paraphrasing ( $\alpha=.76$ ), linking with prior knowledge ( $\alpha=.74$ ), looking at titles and pictures ( $\alpha=.65$ ), study approach ( $\alpha=.60$ ), self-monitoring ( $\alpha=.72$ ), and self-evaluation ( $\alpha=.68$ ).

After completing this questionnaire, *text acquisition* was tested with a free recall test. The recall score represents the percentage of recalled text information. To conclude, pupils were then given back their studied text and were asked to make a graphic organizer or Mind Map of the text, to gain insight in their *schematisation skills*. The shape, organization, and content of the maps are analysed by means of a previously pilot-tested quantitative scoring rubric (Merchie, 2009).

### *Data analysis*

The cognitive load rating scores were analysed by means of Univariate Analyses of Variance. The subscales of the questionnaire and the Mind Map scoring rubric were analysed by means of One-Way Repeated-Measures Analyses of Variance.

## **Results and Conclusion**

The author-provided MM condition perceived significantly less *cognitive load* than the condition working with student-generated MM ( $F(331,1) = 8.196, p=.004$ ). Consequently, working with author-provided MM at first appears more recommendable for this age group than introducing the active construction of MM immediately. However, further research is necessary to confirm these results regarding cognitive load and to study the relation with students' actual learning.

Concerning the evolution in *learning strategies*, a differentiated pattern occurred. Significant effects were mainly found for fifth graders. More particularly, significant differences between the conditions were found for low achievers as to their reported summarising ( $F(53, 2)= 5.787, p=0.005$ ), highlighting ( $F(53,2)=4.978, p=0.001$ ), and self-evaluation strategies ( $F(48,2)=3.411, p=0.042$ ). Interestingly, low achievers in the author-provided MM condition overall score higher on these subscales than in the student-generated condition, whereas the control condition reports to highlight

less, and to summarise and self-evaluate more. Concerning high achievers, significant differences were found for highlighting ( $F(121,2)=3.186$ ,  $p=0.045$ ), rereading ( $F(121,2)=4.400$ ,  $p=0.014$ ), and paraphrasing ( $F(123, 2)=8.989$ ,  $p=0.000$ ). More particularly, the control condition outperforms the experimental conditions on the last two subscales and the student-generated MM condition overall reports lower scores for highlighting. For average achievers, no significant differences were found between the conditions. When interpreting these findings, the self-report nature of the questionnaire must be taken into account and results should therefore be complemented further with trace methodology analyses on the texts and scratch papers.

Concerning *text acquisition*, no significant effects on text recall were found between the conditions ( $F(582,2)= 1.941$ ,  $p=0.144$ ). This result might be due to the fact that the trained strategies were of no direct help for immediate global text recall. Another explanation is that students from the MM conditions did not explicitly apply the trained learning strategies due to the short time span wherein they had to study the relatively short text. The administration of a delayed recall test could have shed another light on this finding (Stull & Mayer, 2007).

With regard to the *schematisation skills*, preliminary analyses illustrate striking differences between the conditions (Figure 1). Whereas the experimental conditions clearly evolve in content organisation and summary, the control condition stays focused on copying text information without any spatial organization. Remarkably, students from the author-provided MM condition who received no explicit training in the construction of MM, also seem to be able to create MM.

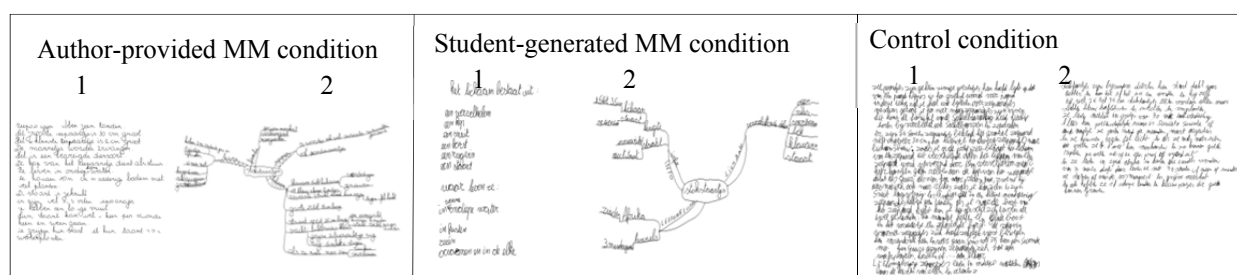


Figure 1. Evolution of schematisation skills within the three conditions from pretest (1) to posttest (2)

## References

- Bakken, J. P., & Whedon, C. K. (2002). Teaching text structure to improve reading comprehension. *Intervention in School & Clinic*, 37(4), 229-233.
- Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: a theoretical inquiry. *Research Papers in Education*, 21(2), 119-132.
- Buzan, T. (1974). *Use your head*. London: BBC Books
- Farrand, P., Hussain, F., & Hennessy, E. (2002). The efficacy of the 'mind map' study technique. *Medical Education*, 36, 426-431.
- Merchie, E. (2009). An intervention study on the use of Mind Maps in the fifth and sixth grade of primary education. Unpublished thesis. Ghent: Ghent University, department of Educational Studies
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429-434.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 106, 1854-1878.
- Rawson, K. A., & Dunlosky, J. (2007). Improving students' self-evaluation of learning for key concepts in textbook materials. *European Journal of Cognitive Psychology*, 19, 559-579.
- Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology*, 99(4), 808-820.
- Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology review*, 14(3), 261-312



# Dynamic Representations on the Interactive Whiteboard

*Jan van der Meij, Hans van der Meij*  
University of Twente  
PO Box 217 7500AE Enschede, The Netherlands  
j.vandermeij@utwente.nl, h.vandermeij@utwente.nl

**Abstract.** In this study we assessed whether presenting dynamic representations on an IWB would lead to better learning gains compared to presenting static representations. Participants were 7-8 year old primary school children learning about views (N = 151) and the water cycle (N = 182). The results showed significant learning gains from pre-test to post-test for both topics in both conditions. Only for the topic views a significant interaction was found between time and condition.

**Keywords:** Static Representations, Dynamic Representations, Interactive Whiteboard

## Introduction

There is a rapid increase of interactive whiteboards (IWBs) in schools. This is especially the case in primary education. In The Netherlands, for example, 98% of the primary schools have at least one IWB and many schools already have IWBs in every classroom (ten Brummelhuis & van Amerongen, 2010). The explosive increase in the number of IWBs is easily explained: the interactive whiteboard is an example of an ICT application that is easy in use. Because the digital whiteboard has many advantages over the traditional chalkboard, their introduction has led to rapid acceptance by teachers, learners and parents. Although both teachers and learners are enthusiastic about using an IWB, Smith, Higgins, Wall, and Miller (2005) and (Lee, 2010) indicate that it is still unclear whether this translates into effective and targeted practice, as there is little research on learning performance.

One unique characteristic of the IWB over the traditional chalkboard is that it can bring dynamic representations into the classroom and that these representations can be manipulated directly (Greiffenhagen, 2000). Dynamic representations can show changes as a function of time (e.g., animations) or can show objects for which manipulation in three dimensions is important (e.g., rotation of objects). In this study we investigated the impact of both types of dynamic representations.

Although dynamic representations are often attractive and are assumed to have added value over static representations, research shows that they do not always lead to better learning (e.g., Boucheix & Schneider, 2009; Tversky, Bauer Morrison, & Betrancourt, 2002). Rieber (1990) recommended that dynamic representations should only be used if they are consistent with the learning task. For use on the IWB this means that dynamic representations should only be used for topics where static representations are insufficient to show the characteristics and properties of a particular topic.

Research on comparing static and dynamic representations so far is mainly done in settings where students were offered representations through self-instruction on a computer. In our study, the teacher has a central role. The teacher uses the representations to explain the topic.

The goal of this study was to assess whether using dynamic representations of (1) views and (2) the water cycle on the IWB lead to better learning outcomes compared to using static representations. We expected that the dynamic conditions would lead to larger learning gains as in this condition the participants would benefit from being able to observe the transitions between the different viewpoints (views) and stages (water cycle).

## Method

### *Participants*

For views, the participants were 151 group 4 children (age 7-8 year) from 9 primary schools. For the water cycle, the participants were 182 group 4 children (age 7-8 year) from 10 primary schools.

## Materials

For the presentation of the IWB materials a website was developed. Figure 1 shows an example for each topic. Each topic was covered in two lessons of 20 minutes. For views, the sequence of the lessons was based on difficulty level: up to 5 blocks in lesson one and up to 10 blocks in lesson two. For the water cycle, the sequence varied in abstraction: heating water in a pan in lesson one and heating seawater by the sun in lesson two.



Figure 1. Website of views (left) and water cycle (right)

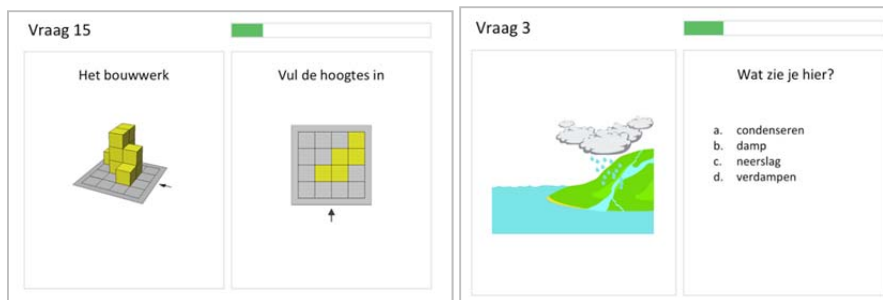


Figure 2. Examples of views (left) and water cycle (right) test items

A teacher manual gave a detailed description of each lesson including all the steps to perform on the IWB, guidance for explaining the topic, and questions for class discussions.

For each topic, a pre-test and post-test consisted of 16 questions. The construction of the test items for views was based on a matrix with the dimensions: context (same / different and more than 5 blocks / up to 10 blocks) and activity (counting cubes, recognizing view, recognizing ground plan and filling in the number of blocks in a ground plan). The construction of the items for the water cycle was based on a matrix with the dimensions: figure (same / different and realistic / schematically), activity (isolated recognizing, recognition, and integration sequence). All tests were presented on the IWB. Figure 2 shows a test item for both topics. Each test item was timed. The timing was based on a pilot. The students gave their answers on paper.

## Results

The average pre-test and post-test scores for the static condition of *views* were 6.07 ( $SD = 2.33$ ) and 8.81 ( $SD = 2.41$ ) respectively. The average pre-test and post-test scores for the dynamic condition were 5.96 ( $SD = 2.35$ ) and 9.75 ( $SD = 2.18$ ) respectively. A repeated measures ANOVA showed significant learning gains for both conditions,  $F(1,149) = 290.12$ ,  $p = 0.00$ , and an interaction effect between time and condition,  $F(1,149) = 7.12$ ,  $p < 0.01$ . Participants in the dynamic condition ( $d = 3.61$ ) showed larger learning gains compared to participants in the static condition ( $d = 2.26$ ).

The average pre-test and post-test scores for the static condition of the *water cycle* were 5.27 ( $SD = 2.40$ ) and 9.43 ( $SD = 2.73$ ) respectively. The average pre-test and post-test scores for the dynamic

condition were 5.70 ( $SD = 2.25$ ) and 9.22 ( $SD = 3.24$ ) respectively. A repeated measures ANOVA showed significant learning gains for both conditions,  $F(1,180) = 236.57$ ,  $p = 0.00$ , but no interaction effect between time and condition,  $F(1,180) = 1.65$ , n.s. Participants in the dynamic condition ( $d = 2.43$ ) did not show larger learning gains compared to participants in the static condition ( $d = 2.16$ ).

### Discussion and conclusion

A special feature of this study is that it examined the effectiveness of dynamic versus static representations in teacher-led lessons with the IWB. In preparing for this study we initially opted for using existing materials, but these were found deficient and so we ended up developing these ourselves. The findings indicated that the lessons led to substantial learning gains on both topics. For views we found the predicted superiority of the dynamic representation, for the water cycle we did not. We will discuss the explanations for this finding during the SIG meeting.

### References

- Boucheix, J.-M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19, 112-127.
- Greiffenhagen, C. (2000). Out of the office into the school: electronic whiteboards for education. *Programming Research Group Technical Report TR-16-00*. Retrieved February, 11, 16-00.
- Lee, M. (2010). Interactive whiteboards and schooling: the context. *Technology, Pedagogy and Education*, 19, 133 - 141.
- Rieber, L. (1990). Animation in computer-based instruction. *Educational Technology Research and Development*, 38, 77-86.
- Smith, H., Higgins, S., Wall, K., & Miller, J. (2005). Interactive whiteboards: boon or bandwagon? A critical review of the literature. *Journal of Computer Assisted Learning*, 21, 91-101.
- ten Brummelhuis, A., & van Amerongen, M. (2010). Vier in Balans Monitor 2010. Ict in het onderwijs: de stand van zaken. Zoetermeer: Kennisnet.
- Tversky, B., Bauer Morrison, J., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.

# Does Targeting the Situation Model during Retrieval Promote Transfer from Expository Texts?

*Stijn van Mierlo, Huib Tabbers, Fred Paas*

Erasmus University Rotterdam

Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands

s.a.a.vanmierlo@fsw.eur.nl, tabbers@eur.nl, paas@fsw.eur.nl

**Abstract.** Retrieving previously acquired information from memory makes that information more robust to forgetting than restudying. This study investigated the effect of two retrieval strategies on transfer from texts. After reading a difficult expository text, participants were allocated to a generative writing condition (describe what you comprehended), a free recall condition or a reread condition. A transfer test, a sentence verification task and an inference verification task were administered immediately or one week after retrieval. From a Construction-Integration Model (Kintsch, 1988; 1998) perspective and based on the retrieval hypothesis (Dempster, 1996) it was expected that participants in the generative writing condition outperformed the participants in the other conditions on a delayed transfer test, but not on a delayed inference or sentence verification task.

**Keywords:** testing effect, transfer, situation model, generative writing, free recall

## Introduction

Retrieving previously acquired information from memory makes that information more robust against forgetting than restudying, a phenomenon referred to as the testing effect. These effects of retrieval practice are typically observed after a delay. It is found to be beneficial on a wide range of materials, in the laboratory, as well as in applied classroom settings (for reviews see Delaney, Verkoeijen, & Spiguel, 2010; Roediger & Karpicke, 2006a). An explanation for this phenomenon can be found in the retrieval hypothesis (Dempster, 1996). This theory states that what is retrieved from memory is less susceptible to decay compared to what is passively restudied.

Whereas much of the past research has focused on delayed memory for text material, from an educational perspective it is more interesting to look at the delayed ability to apply information to a new situation, i.e. transfer. In this experiment we investigated whether generic retrieval strategies are beneficial to delayed transfer performance.

In order to be beneficial for delayed transfer performance, retrieval practice should be congruent with the level of application. The rationale for this can be found in the Construction-Integration Model (C-I Model: Kintsch, 1988; 1998). The C-I Model distinguishes three levels of text representation that an individual concurrently creates from a text. First, the surface level, which contains the surface features of a text. Second, the text base level, which comprises the connection of ideas within a text. Third, the situation model level, containing all inferences that cannot be derived from the text itself, such as connections of information from the material to prior knowledge. The C-I Model states that a coherent situation model leads to higher levels of comprehension and this is what an individual needs to be able to apply acquired information from the text to new situations. So, in order to be beneficial for delayed performance on a transfer task, retrieval practice should be aimed at the consolidation of the situation model of a text.

Support for this congruency principle can be found in the literature. Free recall, which can be considered as a text base retrieval strategy from the C-I model perspective, results in superior performance on delayed memory tests (Glover, 1989; Roediger & Karpicke, 2006b). However, studies have been found to indicate that free recall is not beneficial for delayed transfer performance (Dirkx, Kester, & Kirschner, 2011; Johnson & Mayer, 2009). Testing with conceptual or transfer questions, which can be considered as a situation model based retrieval strategy from the C-I perspective, led to better performance on a delayed transfer test (Butler, 2010; Johnson & Mayer, 2009). To our best notice, this is the first study to investigate a generic retrieval strategy targeting the situation model

within the testing effect design. To manipulate retrieval at the situation model level, participants are asked to type into a computer in their own words what they have comprehended from the text, allowing more room for elaboration and making inferences beyond the text.

The reason to choose for generic retrieval strategies instead of more directed retrieval strategies in this study is because of its opportunities for practical applicability. If the hypothesized effect is found, teachers do not have to hand out possible final test questions or have to generate extra questions for students and students are able to apply these strategies independent of the presence of study questions.

Furthermore, this study differs on an important issue from earlier studies investigating delayed transfer performance. That means, this study does not provide feedback containing correct answers to participants after retrieval (e.g. Butler, 2010). For this reason, the effects of this study can be attributed to retrieval practice, feedback or a combination of both. By providing no feedback, providing no restudy opportunity and using final test questions not used previously in the experiment, it is possible to attribute results from the current study to retrieval practice alone and therefore this study provides better evidence to answer the question what the effect of retrieval is on delayed transfer performance than previous studies.

Following the C-I Model (Kintsch, 1988; 1998) and the retrieval hypothesis (Dempster, 1996), it is hypothesized that retrieval practice focused on the situation model leads to better performance on a delayed transfer test compared to retrieval practice focused on lower levels of the C-I Model.

## **Method**

### *Participants and Design*

Participants were 72 Psychology undergraduate students from Erasmus Universiteit Rotterdam. The design was a 3 (retrieval condition: generative writing, free recall, control) x 2 (time of final test: immediate, delayed) mixed design. Retrieval condition was manipulated between participants, time of final test was manipulated within subjects. Participants were randomly assigned to one of the three retrieval conditions.

### *Materials*

*Texts.* Four expository texts were used, all with a length of 264-299 words. Pilot testing indicated that texts of this length were suited to the manipulation. Texts were presented to participants on a computer screen and in Dutch. Reading order was counterbalanced.

*Comprehension tests.* A short answer transfer test was used to measure deep comprehension. In this test participants had to apply acquired information to a new situation. Because our main interest lies in transfer performance and to prevent interference from the two other comprehension tasks, this test was always administered first. Furthermore, a sentence verification task and an inference verification task were used to measure surface level and text base level understanding respectively.

### *Procedure*

Participants were all individually tested, with the experimenter being present in the same room during the experiment. A practice text and retrieval practice strategy were used to make participants acquainted with the procedure from the experiment. Participants had 5 minutes to study a text and were told they would be tested on the content afterwards. After studying the first text, participants received differentiated instructions. For the restudy condition, instructions were that participants had extra study time before being tested on the material. For the free-recall condition, instructions to participants were that they had to type on a computer everything they could remember from the studied text. For the generative writing, participants were instructed to type a description of what they

learned from the text in their own words. For all tasks, participants had 5 minutes. Participants received the comprehension tasks for two of the texts immediately and for the two remaining texts after one week delay.

## Results

The study is currently conducted. Results will be presented during the EARLI Sig2 meeting in August 2012.

## References

- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1118-1133.
- Delaney, P. F., Verhoeijen, P. P. J. L., Spigler, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 53, pp. 63-148). Burlington: Academic Press.
- Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E.L. Bjork & R.A. Bjork (Eds.), *Human memory* (pp 197-236). San Diego, CA: Academic Press.
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2011, September). *The effects of generative testing on text retention and text comprehension*. Paper presented at the annual meeting of the European Association for Research on Learning and Instruction, Exeter, United Kingdom.
- Glover, J. A. (1989). The “testing” phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, 81, 392–399.
- Johnson, C. I., & Mayer, R.E. (2009). A testing effect with multimedia learning. *Journal of Educational Psychology*, 101, 621–629.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction – integration model. *Psychological Review*, 95, 163–182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Roediger, H. L., III, & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181–210.
- Roediger, H. L., III, & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, 91, 301-311.

# A Simple Scale of Noun Abstractness for Predicting Popular-Scientific Text Comprehension

Jaan Mikk

University of Tartu, Estonia

Jaan.Mikk@ut.ee

**Abstract.** Text abstractness is a well-known but rarely used in practice feature of text comprehensibility. The proposed scale of noun abstractness divides the nouns into three categories: nouns signifying directly perceivable objects, nouns signifying perceivable activities and phenomena, and nouns signifying directly imperceivable notions. The aim of the research was to find the validity of the scale in predicting popular-scientific text comprehension. In three studies with 48, 30 and 40 texts, students learned the texts and filled in the blanks in cloze tests and/or answered the content questions of post-test. The correlations between the average noun abstractness level and the percentage of correct answers in the tests were from  $-.52$  to  $-0.66$ . It would be useful to compose the dictionaries of noun abstractness for computer-based analysis of text comprehensibility.

**Key words:** Text comprehension, level of abstractness, popular-scientific texts

Experimental investigations have shown that abstract material is difficult to understand and remember (Eye, Dixon & Krampen 1989; Mikk, 2000). Sadoski, Goetz and Rodrigues (2000) have found path coefficient 0.81 from text concreteness to text comprehensibility. D. Geshke et al. (2010) have found that more abstract text is used to describe negative behaviour of other groups' members.

There have been made several attempts to measure the level of abstractness of text. We based our simple scale of noun abstractness on the visibility of the objects denominated by the noun. Our three stage scale of noun abstractness is as follows:

- 1 - nouns signifying directly perceivable objects (e. g. ball, dog, car),
- 2 - nouns signifying perceivable activities and phenomena (e. g. run, sunshine),
- 3 - nouns signifying directly imperceivable notions (e. g. cause, function, neutrino).

The reliability of applying the scale was assessed using the test-retest method. A specialist well-trained in using the scale rated the abstractness level of 128 nouns. A year later, he rated the levels once more without looking at the first ratings. The correlation between the first and second ratings was .88.

The aim of the studies below was to evaluate the validity of the noun abstractness scale in predicting popular-scientific text comprehension. We hypothesized that texts with higher level of abstractness of nouns are more difficult to understand and remember.

There have been carried out three studies in a similar way and we will treat them together.

## Method

Overview of the *subjects and texts* in the studies is given in Table 1.

*The abstractness of nouns* in the texts was assessed according to the three stage scale above. We differentiated the assessment of abstractness from the familiarity of nouns. For example, "atom" is a rather frequent and known word, but we assessed it's abstractness 3 because atom cannot be seen. On the other side, our students have not seen pyramids in Egypt but these can be seen in principle and so the noun "pyramid" had abstractness 1.

## Procedure.

Text comprehension was measured by cloze tests and/or post test. In cloze tests, every seventh word in the texts was replaced by a blank and students had to fill in the blanks. Seven versions of cloze tests were composed for every text. Post tests were also composed in several versions for better covering the content of the texts. The questions in tests were aimed at recalling concepts and facts, at

comprehension of connections and phenomena, transforming information and at solving problems. Every student worked with every text without any support from teachers or co-students. Different students worked with different texts in each class. The inter-scorer reliability for scoring the free-response questions in post-test was .85. Detailed description of the method is published elsewhere (Mikk, Kukemelk, 2010).

Table 1: Subjects and texts in the studies.

	I study Popular-scientific texts in biology	II study Popular-scientific texts	III study Texts from physics textbooks
Subjects	124 students 15–16 years old	30 students 15–17 years old	393 students 16–17 years old
Texts	48 texts in Russian Average length 258 words	30 texts in Estonian Average length 157 words	40 texts in Russian Average length 499 words
The level of prior knowledge	5.5%	29.7%	26.1%
Post-test score	36,2%	49.0%	54.4%

## Results and Discussion

Arithmetic mean of text nouns' abstractness was calculated for every text. Nouns with the abstractness level 1 were considered concrete nouns and their percentage among all nouns of the text was also calculated for every text. Abstract nouns were the nouns with the abstractness level 3.

The indices of text comprehension were correlated with the indices of noun abstractness (Table 2).

Table 2: Correlations between text abstractness and comprehension.<sup>6</sup>

Indices of abstractness	I study Cloze procedure	I study Post-test score	II study Cloze procedure	III study Post-test score
Mean abstractness of nouns	–.52	–.66	–.53	–.53
Percentage of concrete nouns	.45	.58	.41	.51
Percentage of abstract nouns	–.64	–.71	–.38	–.40

All the correlation coefficients in the Table 1 indicate that texts with higher portion of abstract nouns are more difficult to understand. The more there are concrete nouns in the text the easier is the text. The validity of the average noun abstractness was from 0,52 to 0,66 in predicting reading comprehension.

In the study of physics textbooks the correlations were lower than in the study popular scientific texts of biology and in the same time, the prior knowledge of section content in physics textbooks was higher than in the case of popular scientific texts in biology.

Abstract nouns are longer than concrete words. J. Elts (1995) has found correlation 0.96 between the abstractness level of nouns in biology texts and the length of nouns in letters. Despite that the average level of noun abstractness was a more powerful predictor of text readability according to Estonian readability formula for popular scientific texts (Mikk, 2000, 118) and for Russian texts in physics (Kukemelk & Mikk, 1993, 100).

Considering the importance of text analysis by computers, it would be useful to compose dictionaries of noun abstractness. Some examples of such dictionaries are available (Della Rosa, 2010; Guido, 2004).

<sup>6</sup> All the correlation coefficients are statistically significant at .05 level.



## Limitation

The samples of texts and students were not representative to popular-scientific texts or to all students in Estonia. Therefore, we cannot generalize the results, however, the concurrency of the results supports the validity of the scale in the wide area of texts and students.

## References

- Della Rosa, P. A., & Catricalà, E., & Vigliocco, G., & Cappa S. F. (2010). Beyond the abstract–concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods*, 42(4), 1042–1048.
- Elts J. (1995). Word length and its semantic complexity. In I. Kraav, J. Mikk, L. Vassiltchenko (Eds.) *Family and Textbooks*. Tartu: University of Tartu, 115–126.
- Eye, A., & Dixon, R. A., & Krampen, G. (1989). Text recall in adulthood: The roles of text imagery and orienting. *Psychological Research*, 51, 136–146.
- Geschke, D., & Sassenberg, K., & Ruhrmann, G., & Sommer, D. (2010). Effects of linguistic abstractness in the mass media: How newspaper articles shape readers' attitudes toward migrants. *Journal of Media Psychology: Theories, Methods, and Applications*, 22(3), 99–104.
- Guido, G., & Provenzano, M. R. (2004). Abstractness and emotionality values for 398 English words. *Perceptual and Motor Skills*, 98(3), 1265–1268.
- Kukemelk, H., Mikk, J. (1993). The prognosticating effectivity of learning a text in physics. *Glottometrica*, vol. 14, 82–103.
- Mikk, J. (2000). *Textbook: Research and Writing*. Frankfurt am Main etc.: Peter Lang.
- Mikk, J. & Kukemelk, H. (2010). The relationship of text features to the level of interest in science texts. *TRAMES: Journal of the Humanities and Social Sciences*, 2010, 14(64/59), 1, 54 – 70.
- Sadoski, M., & Goetz, E. T., & Rodriguez, M. (2000). Engaging texts: Effects of concreteness on comprehensibility, interest, and recall in four text types. *Journal of Educational Psychology*, 92, 85–95.

# Meaningful Rereading for Meaning: the Role of Self-Regulation in Successful Reading Procedures

Marta Minguela, Isabel Solé

Department of Developmental and Educational Psychology. University of Barcelona.

Pg. de la Vall d'Hebrón, 171. 08035, Barcelona, Spain

*martaminguela@ub.edu, isoleg@ub.edu*

**Abstract.** The main goals of this study were to describe the reading procedures used by last-year secondary school students; to investigate whether skilled and less-skilled readers use different reading procedures; and to find out the relation between these procedures and reading comprehension. The reading process followed by 55 secondary school students was examined, keeping track of their procedure using specific software. Participants read an expository text and answered comprehension questions afterwards. Four reading procedures were identified, differing in terms of self-regulation. Skilled readers tended to use non-linear and recursive patterns, while the more mechanical patterns were used by less-skilled readers. Significant differences were found in high level comprehension performance (but not in low level questions) depending on the reading procedure used; these differences could be attributed neither to reading time nor to general comprehension ability. Those using more self-regulated reading procedures performed better in high level questions than those reading mechanically.

**Keywords:** reading comprehension, reading procedures, self-regulation, monitoring comprehension

## Introduction

Being fully competent in reading implies not only superficial comprehension (involving low level processes, i.e. processing literal information or doing simple inferences), but also deep comprehension (demanding high level processing, like integrating information across the text, or reasoning beyond text) (Graesser, Singer & Trabasso, 1994). Skilled readers, faced with different requirements, know which cognitive processes need to be activated and how to do it; i.e., they can self-regulate comprehension. This self-regulation is reflected in the product of comprehension, as well as in the procedures in which a reader engages. Prior research has described different reading procedures (Goldman & Saul, 1990), suggesting that skilled and less-skilled comprehenders differ in their involvement in metacognitive strategies (Garner & Reis, 1981). Successful self-regulation would enable readers to assess the reading process ("monitoring") and to adjust their behaviour to the reading aim or to incomprehension ("control"), i.e. to use adequately the two main components of metacognition (Nelson, 1996). One of the possible control strategies is rereading, commonly assumed to be adequate to fix incomprehension or to deepen comprehension, and frequently considered in the study of reading procedures, as it implies recursion, in comparison to linear procedures. The use of rereading must imply some kind of reading "monitoring". So, it could be hypothesized that its profitable use may be related to self-regulation. However, some approaches to the study of rereading neglect the importance of self-regulation, automatically assigning participants to one condition or another (typically a "read the text once" vs. a "read the text twice" condition, e.g. Callender & McDaniel, 2009). In order to try to remove this limitation, the present study aimed at describing the reading procedures used by students when they read following any procedure they want; at investigating whether these procedures differ between skilled and less-skilled readers; and at finding out if different reading procedures lead to different comprehension results.

## Method

Participants were 55 final-year secondary school students attending three similar schools in Barcelona, Spain. They were selected considering two variables: *prior knowledge* (that has to be low) and *reading comprehension ability* (measured with a standardized test (TEC, Vidal-Abarca et al., 2007). We selected participants scoring above percentile 80 (*skilled readers*) and those situated in 30th-40th percentile (*less-skilled readers*); readers scoring in 20<sup>th</sup> percentile or less were excluded). As data were

collected using specific software to keep track of the reading process (Read&Answer, Vidal-Abarca et al., 2011), we used sample material to instruct participants in its workings, and to illustrate the kind of task that they will be required to do later. Once they were ready, they went on to the critical trial. Then, participants had to read an expository text about “The Second World War” (676 words), directly extracted from a Social Sciences textbook. They were told to read it carefully, considering that they would have to answer questions about it later. After reading, they answered to these questions. The reading comprehension questions were constructed *ad hoc* for this experiment, and consisted on 4 low level questions (superficial comprehension) and 5 high level questions (deep comprehension). Participants performed the task individually, in the presence of a researcher, and with no time limit.

## Results and Discussion

Reading procedures were analysed considering reading sequences (to establish if the reading was linear or introduced some kind of recursion during or after the process) and reading speed for each segment of the text (*stability envelope* analysis (Gast & Spriggs, 2009) were performed to determine to what extent reading speed has been stable across the different segments of the text). Consistent with Goldman & Saul (1990), four different reading procedures were found: *linear reading* (a unique sequential reading at a stable speed), *linear rereading* (two sequential readings; the first reading is slower than the second one, but both are done at a balanced pace), *linear reading followed by selective rereading* (a sequential reading with rereading of some fragments afterwards; reading speed is more variable), and *reading with online rereading* (incorporates some punctual rereading before reaching the end of the text, and reading speed is different across segments).

Figure 1 shows the percentage of participants using each procedure, considering the whole sample (N=55). Figure 2 represents this information, differentiating amongst skilled (N=29) and less-skilled readers (N=26).

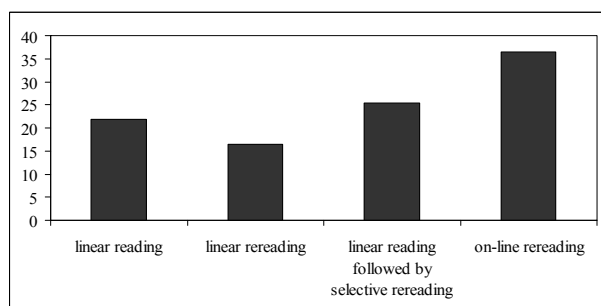


Figure 1. Percentage of reading procedure use amongst skilled and less-skilled readers.

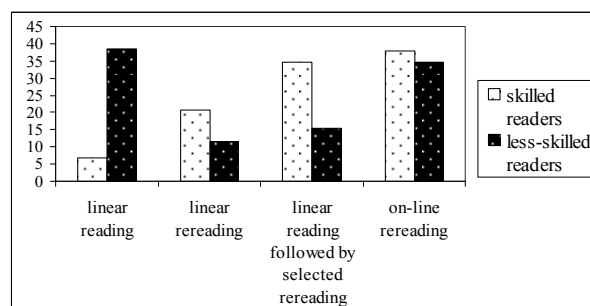


Figure 2. Percentage of reading procedure use amongst skilled and less-skilled readers.

*Linear reading* procedure was essentially used by less-skilled readers, whilst the *linear reading followed by selected rereading* procedure is more used amongst skilled readers. It could be argued that this procedure incorporates self-regulation to some extent, because participants reread only some segments of the text and devote different attention to different parts of the text (opposite to a mechanical reading). The same could be said about the *reading with on-line rereading* procedure, which turned out to be the most commonly used procedure in our sample, and curiously enough, both skilled and less-skilled readers used it in a similar proportion.

To investigate whether participants obtained different comprehension results as a function of the reading pattern used, ANOVA analyses were conducted separately for each reading comprehension subscale. For low level questions, results showed no significant main effect of the reading procedure  $F(3, 51) = 1.846, p = .151, \eta_p^2 = .098$ . Conversely, for high level questions, a main effect of the reading procedure used was found  $F(3, 51) = 4.59, p = .006, \eta_p^2 = .213$ . In view of these results,

ANCOVA analyses were performed for these questions, to examine if the significant main effect obtained could be attributable to the reading procedure, or to other covariables such as reading comprehension ability and time spent on reading the text. Results showed that, once we controlled the effect of these two covariates, the main effect of the reading procedure still was significant,  $F(3, 49) = 3.188$ ,  $p = .032$ ,  $\eta_p^2 = .163$ . Specifically, those using a *linear reading followed by selected rereading* and a *reading with on-line rereading* procedure obtained better results than those reading and rereading the text linearly.

So, these first two procedures seemed to be better self-regulated and more *strategic*, considering that, in fact, led to better comprehension performance. These results are consistent with the idea that, when having low prior knowledge, one linear reading is not enough; at the most, it may be useful to superficially comprehend a text, but it would difficultly enable deep comprehension (Sánchez, 1998). Also, results suggest that when it comes to deep comprehension, reading procedures and rereading strategies are more effective if the reader adapts its use to the ongoing comprehension, than when they are used mechanically. These results seem to point out the relation between comprehension and self-regulation: skilled readers tend to use more self-regulated procedures, which are especially important to deeply comprehend texts (but not to process texts superficially). So, one critical issue to achieve high level comprehension may be the way one self-regulates the own reading process.

## References

- Callender, A. A. & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, 34, 30-41.
- Garner, R. & Reis, R. (1981). Monitoring and resolving comprehension obstacles: an investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *RRQ*, 18, 439-447.
- Gast, D. L., & Spriggs, A. D. (2009). Visual analysis of graphic data. In D. L. Gast (Ed.), *Single subject research methodology in behavioral sciences* (pp. 199–233). London: Routledge.
- Goldman, S. R. & Saul, E. U. (1990). Flexibility in text processing: a strategy competition model. *Learning and individual differences*, 2, 181-219.
- Graesser, A.C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371-395.
- Nelson, T.O. (1996). Consciousness and metacognition. *American Psychologist*, 51, 102-116.
- Sánchez, E. (1998). *Comprensión y redacción de textos*. Barcelona: Edebé.
- Vidal-Abarca, E., Gilabert, R., Martínez, T. & Sellés, P. (2007). *Test de Estrategias de Comprensión*. Madrid: ICCE.
- Vidal-Abarca, E., Martínez, T., Salmerón, L., Cerdán, R., Gilabert, R., Gil, L., Mañá, A., Llorens, A. & Ferris, R. (2011). Recording online processes in task-oriented reading with Read&Answer. *Behavior Research Methods Instruments & Computers*. 43, 179-192.

# Animations Facilitate Spatial Perspective Taking

Stefan Münzer

Universität Mannheim

A 5, 68131 Mannheim, Germany

stefan.muenzer@uni-mannheim.de

**Abstract.** Spatial perspective taking means to imagine a spatial configuration from another point of view. The present study investigated whether focused animations can facilitate spatial perspective taking. The experimental task demanded to estimate directions to memorized objects in a spatial scene from an imagined position and orientation within the scene. Static pictures which required imagined reorientation of the self were compared to animations showing the reorientation externally. Individual differences in perspective taking ability were considered. Results showed a large effect in favor of animations for reaction times. An aptitude-treatment-interaction was found for accuracy: Perspective taking ability played a much stronger role in direction estimation accuracy with static pictures than with animations. It is concluded that focused animations can facilitate spatial perspective taking.

**Keywords:** animation, individual differences, spatial ability, perspective taking, aptitude-treatment-interaction

## Spatial Perspective Taking

In many domains (geography, virtual environments, anatomy, mechanical systems, chemistry, etc.), the understanding of the spatial configuration of components of a system or the spatial relations between objects in a scene is an important learning goal. The popular use of animations and interactive 3D-visualization software suggests that those visualizations are considered supportive.

Static pictorial representations of spatial configurations are typically studied with a particular orientation (e.g., maps are studied with north on top). Judgments of relative directions from memory are more difficult if they require reorientation (e.g., Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Rossano & Moak, 1998). By showing the change of the viewpoint externally, animations may facilitate perspective taking because visual-spatial processing is supported.

In the present study, individual differences in spatial perspective taking ability are considered. Perspective taking is associated with a visual-spatial ability factor termed spatial orientation. This factor is contrasted with the spatial visualization factor, because the egocentric viewpoint changes only in orientation tasks, but not in visualization tasks. The distinction has been supported with a new test of perspective taking (Kozhevnikov & Hegarty, 2001; Hegarty & Waller, 2004).

## Experiment

### Method

*Participants.* Fifty-four students of Saarland University (27 females,  $M = 25.0$  years,  $SD = 5.6$  years) took part. Three participants (2 females) were excluded because they missed more than two items of the perspective taking test. For 19 participants out of 51 participants, reaction time data were not recorded due to a technical error.

*Materials.* Each item of the experimental materials showed a virtual room from a bird's eye view which contained a blue and a red post and an indication of a person with a position and an orientation. Participants had to estimate the relative direction to one of the two posts from memory (i.e. after studying the scene). Participants responded by clicking on a circle imagining that "up" corresponded to their heading if they would stand in the scene with the indicated orientation. In the static presentation condition, perspective taking had to be mentally performed. In the animated presentation condition, perspective taking was shown externally (the animation showed a camera flight into the position in the spatial scene). The angle difference in degrees between the initial bird's-eye view and the orientation of the indicated person was experimentally varied (50°, 90°, 130°, 170°). The angles

between the indicated person's position and orientation and the target post were varied over a wide range (e.g., 35°, 70°, 110°, 130°, 150°, 200°, 235°, 295°). Participants performed all conditions.

Perspective taking ability was measured with the perspective taking test (Hegarty & Waller, 2004). A map is presented which shows a spatial configuration of seven objects. Participants imagine themselves standing at a particular position, facing a particular second location, pointing to a third location. The score of the participant is the average angular error over twelve items. The test is administered with a time limit of five minutes. Participants performed the perspective taking test with an average angular error of  $M = 19.70^\circ$ ,  $SD = 11.81^\circ$ ,  $Min = 6.3^\circ$ ,  $Max = 64.6^\circ$  ( $N = 51$ ).

## Results

Descriptive accuracy data (angular error) are shown in Table 1. An analysis of variance with presentation (static, animation) and angle (50°, 90°, 130°, 170°) as within-factors did not reveal an effect of presentation,  $F < 1$ , *ns*, but it yielded a significant effect of angle,  $F(3,156) = 11.768$ ,  $p < .001$ , partial  $\eta^2 = .19$ . An interaction between presentation and angle was not found,  $F < 1$ , *ns*. The 130° and 170° angles were more difficult than the 50° and 90° angles. A Scheffé post hoc test revealed a critical difference of 3.60° ( $p < .05$ ).

Descriptive reaction time data are also shown in Table 1. An analysis of variance with presentation (static, animation) and angle (50°, 90°, 130°, 170°) as within-factors was performed. A significant effect of presentation was obtained,  $F(1,31) = 17.749$ ,  $p < .001$ , partial  $\eta^2 = .36$ . Neither an effect of angle,  $F(3,93) = 1.243$ , *ns*, nor an interaction between presentation and angle was found,  $F < 1$ , *ns*.

Table 1: Accuracy and reaction times separated by presentation and angle.

Presentation condition	Angle condition	Accuracy (angular error)		Reaction times (milliseconds)	
		M	(SD)	M	(SD)
Static	50°	13.49°	(5.99°)	2059.88	(527.41)
	90°	13.68°	(10.86°)	2120.63	(490.59)
	130°	18.29°	(9.66°)	2031.33	(414.62)
	170°	18.05°	(13.26°)	2101.79	(529.32)
Animation	50°	14.14°	(5.59°)	1900.78	(410.36)
	90°	12.95°	(6.37°)	1856.83	(380.24)
	130°	17.60°	(6.02°)	1811.59	(335.53)
	170°	18.35°	(9.07°)	1853.16	(280.66)

Note.  $N = 51$  for accuracy data,  $N = 32$  for reaction time data

An aptitude-treatment-interaction between spatial perspective taking ability and static vs. animated presentation was evaluated by comparing the regression slopes between the static presentation and the animation conditions separately for small (50°, 90° combined) and for large (130°, 170° combined) angles. The comparison of the regression slopes followed Judd, McClelland, and Smith (1996, p. 369). For small angles, the regression coefficient with static presentation was  $r = .57$  ( $r^2 = .33$ ),  $F(1,15) = 24.047$ ,  $p < .001$ . The regression coefficient with animation was  $r = .34$  ( $r^2 = .12$ ),  $F(1,15) = 6.368$ ,  $p < .05$ . The coefficient for the difference slope was  $r = .39$  ( $r^2 = .15$ ),  $F(1,15) = 8.802$ ,  $p < .01$ , indicating a significant difference between the slopes for static presentation vs. animation.

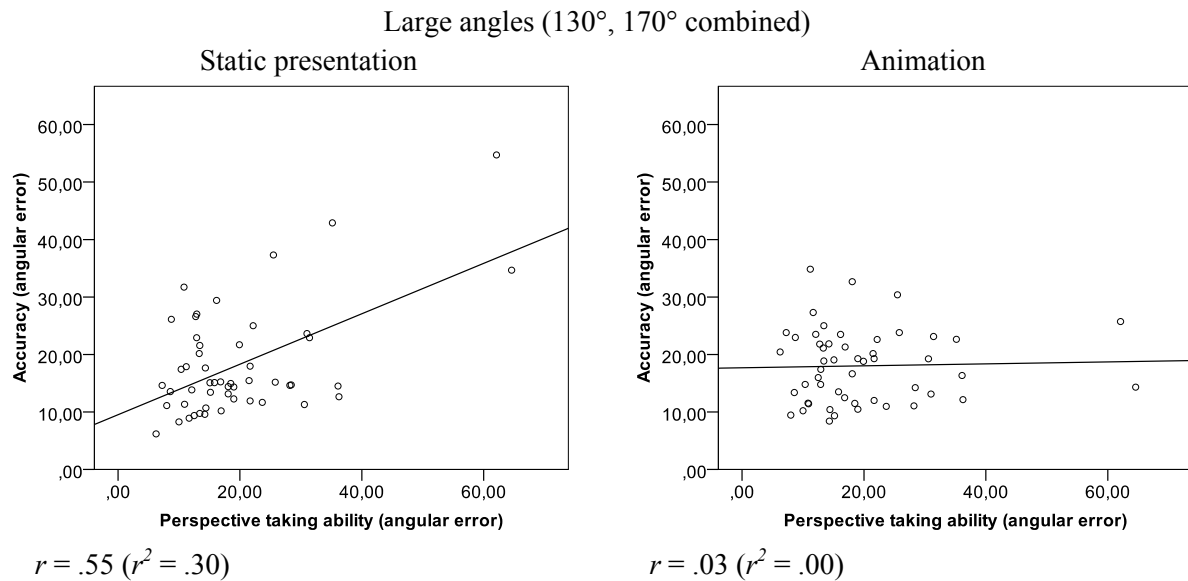


Figure 1. Accuracy (angular error) in the experimental task predicted by perspective taking ability, separately shown for static vs. animated presentation.

For large angles (Figure 1), the regression coefficient with static presentation was  $r = .55$  ( $r^2 = .30$ ),  $F(1,15) = 21.245$ ,  $p < .001$ . With animation it was  $r = .03$  ( $r^2 = .00$ ),  $F < 1$ , *ns*. The coefficient for the difference slope was  $r = .52$  ( $r^2 = .27$ ),  $F(1,15) = 17.889$ ,  $p < .001$ , i.e., a significance difference between the slopes for static presentation vs. animation was found.

## Conclusions

Larger angles between the observers' original position and the to-be-adopted position and orientation resulted in higher angular errors. Thus, difficulty depended on angle. Difficulty was not affected by type of presentation. However, animations allowed participants to answer substantially faster, indicating facilitation. Regression analyses suggest an aptitude-treatment-interaction: The relation between perspective taking ability and accuracy in the experimental task was moderated by type of presentation. Perspective taking ability explained much less variance with animated presentation.

## References

- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking abilities. *Intelligence*, 32, 175-191.
- Judd, C. M., McClelland, G. H., & Smith, E. R. (1996). Testing treatment by covariate interactions when treatment varies within subjects. *Psychological Methods*, 1(4), 366-378.
- Kozhevnikov, M., & Hegarty, M. (2001) A dissociation between object-manipulation spatial ability and spatial orientation abilities. *Memory and Cognition*, 29, 745-756.
- Roskos-Ewoldsen, B., McNamara, T. P., Shelton, A. L., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 215-226.
- Rossano, M. J., & Moak, J. (1998). Spatial representations acquired from computer models: Cognitive load, orientation specificity and the acquisition of survey knowledge. *British Journal of Psychology*, 89, 481-497.

# Modeling the Relationship between Representational Competence and Domain Knowledge

*Sandra Nitz*

Leibniz-Institute for Science and Mathematics Education  
Olshausenstr. 62, 24118 Kiel, Germany  
sandra.nitz@hu-berlin.de

*Claudia Nerdel*

TUM School of Education  
Schellingstraße 33, 8079 Munich, Germany  
nerdel@tum.de

*Helmut Prechtl*

Leibniz-Institute for Science and Mathematics Education  
Olshausenstr. 62, 24118 Kiel, Germany  
prechtl@uni-potsdam.de

**Abstract.** Skills for dealing with external representations in science have been referred to as representational competence. Representational competence enables students to think, communicate, and conceptualise about science concepts and is therefore crucial for developing scientific literacy. Research indicates that students' representational competence is further closely related to students' conceptual understanding. However, previous studies did not empirically distinguish between domain knowledge and representational competence. The aim of this study is therefore to model the development of representational competence and domain knowledge and the relationship between these two constructs. Confirmatory factor analyses and a latent cross-lagged panel model of data from 931 German secondary school students indicate that domain knowledge and representational competence are separable but interconnected constructs and prior knowledge in the domain plays a significant role for the development of representational competence.

**Keywords:** Representational competence, domain knowledge

## Introduction

There is a strong consensus among researchers that learning science involves different multimodal representations such as graphs, diagrams, texts, symbols, and formulae (Ainsworth, 1999, 2006; Kress, Jewitt, Ogborn, & Tsatsarelis, 2001; Lemke, 2004; Yore & Hand, 2010). Interpreting, constructing, transforming, evaluating, and using different scientific representations are therefore crucial skills for students. These skills have been referred to as representational competence which allows students „to reflectively use a variety of representations [...], singly and together, to think about, communicate, and act on [...] phenomena in terms of underlying, a perceptual physical entities and processes” (Kozma & Russell, 2005, 131). Representational competence is situated within a specific domain addressing disciplinary skills that students should acquire to become proficient and literate in the targeted domain and includes the following skills: (a) use representations for describing scientific phenomena; (b) construct and/or select a representation and explain its appropriateness for a specific purpose; (c) identify, describe, and analyze features of representations; (d) compare and contrast different representations and their information content; (e) connect across different representations, map features of one type of representation onto those of another, and explain the relationship between them; (f) take the epistemological position that representations correspond to phenomena but are distinct from them; and (g) use representations in social discourse to support claims, draw inferences, and make predictions (Kozma & Russell, 2005).

Previous research indicates that students' representational competence is closely related to students' conceptual understanding of the domain (Kozma & Russell, 1997; Stieff, 2011). However, these studies did not seek to empirically distinguish between domain knowledge and representational competence. The aim of this study is therefore to model the development of representational competence and domain knowledge in a specific topic and the relationship between these two



constructs. The following research questions are addressed: (1) Are representational competence and domain knowledge separable constructs? (2) What kind of role does students' domain knowledge play in relation to students' representational competence and vice versa?

## Method

To model the development of representational competence and domain knowledge and its relationship we used a one group pretest-posttest design and assessed domain knowledge and representational competence before and after a teaching unit on photosynthesis (15 lessons on average) in 45 grade 11 and 12 classes (German secondary school). In total, 931 students (61% female, 39% male) aged 16 to 24 years ( $M = 17.8$ ,  $SD = 0.7$ ) participated in this study. Confirmatory factor analyses (using Mplus) were used to address research question 1. To address research question 2 a latent cross-lagged panel model was analysed using Mplus.

Instruments with multiple choice items to assess representational competence and domain knowledge (on photosynthesis) were developed and tested in pilot studies. As representational competence is a complicated and broad concept, setting a focus on single aspects was needed for developing the test. The test for representational competence (15 items) focused on the transformation of information among different modes of representations and included the skills a, b, c, and e as described above. Subject-specific information that was necessary to solve the task was given in the item stem to ensure that items could potentially be solved independently from knowledge about photosynthesis. The test on domain knowledge (6 items) focused on central concepts within the topic photosynthesis which are covered in the state-prescribed curriculum for that specific teaching unit (Basel, 2009). Both tests showed acceptable reliability (Cronbachs  $\alpha = .62$  for domain knowledge,  $.68$  for representational competence) and strong measurement invariance.

## Results

Results of confirmatory factor analyses are shown in Table 1. At both measurement times, the model with separated latent factors for domain knowledge and representational competence (two-dimensional) showed a better fit to the data than an alternative model with only one latent factor.

Table 1: Results of Confirmatory Factor Analyses.

	Pre-tests	Post-tests
One-dimensional model	$\chi^2(df) = 228.6 (189)$ , $p=0.026$ ; RMSEA=0.015, CFI/TLI=0.860/0.844	$\chi^2(df) = 254.1 (189)$ , $p=0.001$ ; RMSEA=0.022, CFI/TLI=0.911/0.902
Two-dimensional model	$\chi^2(df) = 218.7 (188)$ , $p=0.062$ ; RMSEA=0.013, CFI/TLI=0.892/0.879	$\chi^2(df) = 224.3 (188)$ , $p=0.036$ ; RMSEA=0.017, CFI/TLI=0.951/0.945

Figure 1 presents the results of the latent cross-lagged panel model which showed good fit indices ( $\chi^2(df) = 876.2 (811)$ ,  $p=0.06$ ; RMSEA = 0.009, CFI/TLI = 0.93/0.92). Domain knowledge and representational competence correlated moderately and share about 30% of variance. The autoregressive regression coefficients indicate that the stability of representational competence from pre- to post-test was more pronounced than the stability of domain knowledge. The asymmetrical cross-lags show a directional effect: Prior domain knowledge positively predicted representational competence but prior representational competence did not predict domain knowledge.

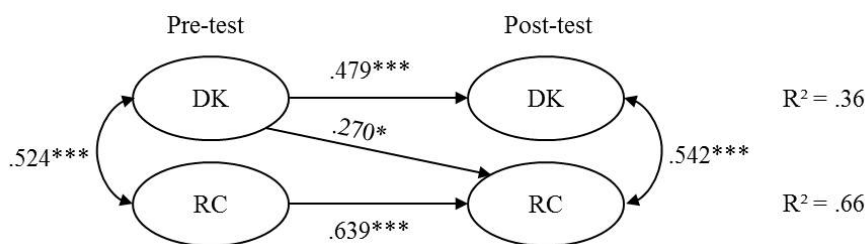


Figure 1. Standardized parameter estimates for the latent cross-lagged panel model (Factors loadings were hold equal across pre-and post-test for each variable, item residuals were allowed to be auto-correlated and only significant parameters are shown. DK = domain knowledge, RC = representational competence).

## Discussion

The analyses indicate that domain knowledge and representational competence are separable but interconnected constructs which is in line with theoretical considerations. Representational competence was rather stable in the analysed time interval of about 15 lessons. Therefore, it might be fruitful to investigate the development over a longer time. Prior knowledge in the domain seems to play a significant role for the development of representational competence. In further studies it is thus necessary to control for domain knowledge when investigating representational competence.

## References

- Ainsworth, S. E. (1999). The functions of multiple representations. *Computers and Education*, 33(2-3), 131–152.
- Ainsworth, S. E. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198.
- Basel, N. (2009). *Entwicklung eines Fragebogens zur Untersuchung des Fachwissens über Stoffwechselphysiologie*. Unveröffentlichte Staatsexamensarbeit, CAU, Kiel.
- Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, 34(9), 949–968.
- Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualizations in science education* (pp. 121–146). Dordrecht: Springer.
- Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (Eds.). (2001). *Advances in Applied Linguistics. Multimodal teaching and learning: Rhetorics of the science classroom*. London: Continuum.
- Lemke, J. L. (2004). The literacies of science. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (pp. 33–47). Arlington: IRA.
- Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *Journal of Research in Science Teaching*, 48(10), 1137–1158.
- Yore, L. D., & Hand, B. (2010). Epilogue: Plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Research in Science Education*, 40(1), 93–101.

# Measuring Representational Competence in Science

*Sandra Nitz*

Humboldt-University Berlin  
Invalidenstr. 42, 10115 Berlin, Germany  
sandra.nitz@hu-berlin.de

*Christine D. Tippet*

University of Victoria  
Victoria, BC, Canada  
chris.tee@shaw.ca

**Abstract.** Science is multimodal - both scientists and students learning science use representations including text, diagrams, graphs, formulas, and symbols to construct knowledge and communicate findings. RC is a recent focus in science education research but there is no widely accepted measure of RC. The projects described here were designed and implemented in different countries (Germany, Canada), grade levels (Grades 11 and 12, Grades 6 and 7), and dealt with different topics (photosynthesis, general science topics). Yet each project attempted to measure RC by asking students to complete tasks that involved constructing representations, choosing the most appropriate representation of a particular concept, analyzing features of representations, and making connections between representations. Each project provided insights into students' interpretation, construction, and transformation of representations and provided a foundation for developing a measure of RC.

**Keywords:** representational competence, assessment

Learning science involves materials that contain multimodal representations (Jewitt, Kress, Ogborn, & Tsatsarelis, 2001). Hence, students need to develop representational competence in order to interpret, construct, relate, and transform representations in a scientific manner. This collection of skills has been referred to as representational competence (RC, Kozma & Russell, 2005). RC allows students to think, communicate, and conceptualize about science concepts and includes the abilities to: (a) use representations for describing scientific concepts; (b) construct and/or select a representation and explain its appropriateness for a specific purpose; (c) use words to identify, describe, and analyze features of representations; (d) compare and contrast different representations and their information content; (e) connect across different representations and explain the relationship between them; (f) realize that representations correspond to phenomena but are distinct from them; and (g) use representations in discourse to support claims, draw inferences, and make predictions.

## The Issue

Given the importance of RC in doing and learning science, it is important to foster those skills in the science classroom. To develop learning environments that foster students' RC, it is necessary to assess students' RC in a comparable manner. Therefore, a measure of RC that is applicable across learning contexts is needed. Experts have noted the lack of such a measure (deVries & Lowe, 2010). Our poster showcases two projects that developed measures for aspects of RC in different contexts utilizing different approaches. We point out advantages and disadvantages of each measure and merge our results to prepare a foundation for developing a reliable measure for assessing RC in science.

## Procedure

The two projects described here were designed and implemented in different countries, at different grade levels, and dealt with different science topics. Yet each project attempted to measure RC with tasks that involved using representations to describe concepts, constructing representations, choosing the most appropriate representation of a particular concept, identifying, describing, and analyzing features or components of representations, and making connections between representations.

Research Project 1 explored how representations are used in biology instruction on photosynthesis in German secondary schools and how instructional approaches to representations affect students' RC. The measure of RC focused on the transformation of scientific information between different representational modes and incorporated skills a, b, c, and e as described above (Kozma & Russell, 2005). 27 tasks with 45 open-ended items were developed based on an analysis of representations in teaching material. The open-ended items were administered in five Grade 11 and 12 classes to explore students' ideas of how to represent a specific topic. Responses were analyzed to identify common errors and closed items were created and administered in three Grade 11 classes to explore the difficulty of these items. Based on these results and theoretical assumptions, 12 tasks with 18 closed items were chosen to be utilized in the final measure (see Figure 1). The final measure was administered in 45 Grade 11 and 12 classes. The unidimensional Rasch model was applied to data from 202 students to obtain information about the items' psychometric quality. Item fit was assessed using the weighted mean square statistic (WMNSQ). WMNSQ of the 18 closed items range from 0.87 to 1.16 showing good item fits. Item difficulties ranged from -1.461 to 1.766. Furthermore, items showed satisfactory discrimination with values between 0.31 and 0.52. The EAP/PV reliability of the test was 0.78, which is considered to be satisfactory. These analyses indicate that the administered items showed satisfactory psychometric quality.

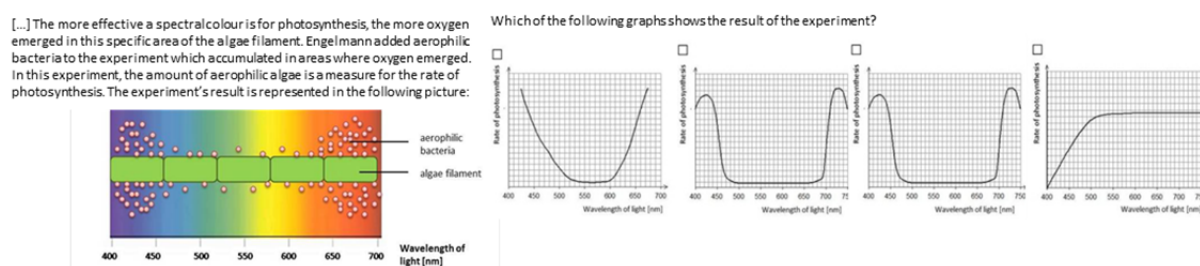


Figure 1. Example of a multiple choice item from Research Project 1. Note: Text is shortened.

Research Project 2 examined Canadian middle school students' RC in science using a variety of researcher-created assessments. Four of the RC skills listed above (Kozma & Russell, 2005) were targeted as appropriate for middle school students: b, c, d, and e. Test items were researcher developed and refined based on feedback from experts in the field and teachers with extensive middle school experience. Items included constructing representations based on unfamiliar science concepts described in a text passage; answering questions based on verbal and visual information and identifying where the answers were located; answering questions that relied on interpretation of spatial diagrams; describing the thought process behind some of the decisions that were made when creating representations; selecting the most appropriate representation for a particular science concept and explaining why it was the most appropriate; and analyzing specific aspects of a representation such as how to depict motion or what an arrow meant (see Figure 2). Tasks were presented in four separate tests to 44 students from four Grades 6 and 7 classrooms and items were assessed for accuracy or content as appropriate. Follow-up interviews were conducted to collect more information about student thinking and interview responses were analyzed iteratively, using open coding. Combined results from tasks and interviews suggest that while many participants could select the most appropriate representations for scientific information, such as flowcharts for processes or cyclical graphics for life cycles, they had difficulty creating their own accurate representations of information, and most participants were unable to articulate their knowledge of representational functions.

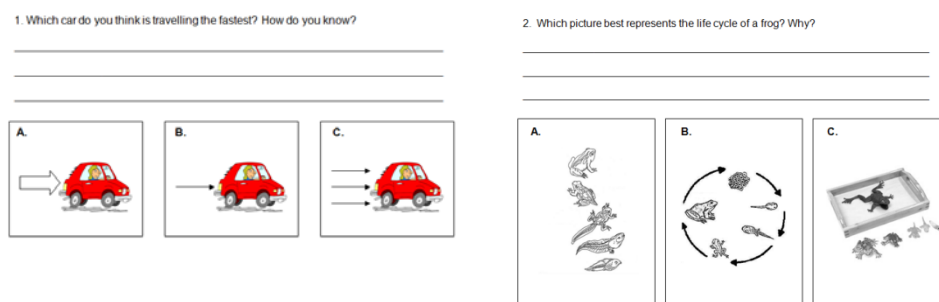


Figure 2. Examples of assessment items from Research Project 2.

### Implications and Contributions

The results of each project have separate implications dependent upon the context. In our separate attempts to design a means of assessing RC, we each adopted a working definition that emphasized a number of discrete skills that needed to be used in an integrated fashion. Our measures included items based on using representations to describe concepts, constructing representations, choosing the most appropriate representation of a particular concept, identifying, describing, and analyzing features or components of representation, and making connections between representations. Although both measures showed satisfactory validity and reliability, further steps need to be made to develop an overall measure of RC. First, we need to rethink the skills described by Kozma and Russell (2005). Do those skills describe RC completely or do other skills need to be addressed as well? The concept of metarepresentational competence (diSessa, 2004) could be a fruitful addition to our consideration of RC. Second, we need to widen our focus and develop tasks and items to cover more skills than those we already incorporated in our respective measures. Third, all items need to be analyzed in terms of their psychometric quality. In this regard we need to think about task and item formats. In our respective projects we used different approaches—Project 1 focused on closed items and a quantitative approach to assess students' RC whereas Project 2 used open items and a qualitative approach. Advantages of both approaches should be incorporated in a measure of RC and IRT models could be used to assess the psychometric quality. These two projects, and the implications yielded by the integration of the projects, contribute to the teaching and learning of science in a number of ways. Each individual project provided insights into how students interpreted, constructed, and transformed representations. Furthermore, we are beginning to develop the foundation for a measure of RC that is applicable in different teaching and learning contexts.

### References

- de Vries, E., & Lowe, R. (2010). *Graphicacy: What Does the Learner Bring to a Graphic?* Paper presented at the Earli SIG 2: Comprehension of text and graphics. Tracing the mind: How do we learn from text and graphics?, Tübingen, Germany.
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22, 293-331.
- Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review*, 53(1), 5-18.
- Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J.K. Gilbert (Ed.), *Visualizations in Science Education* (pp. 121-146). Dordrecht: Springer.

# Implicit and Explicit Guidance for Collaborative Learning at the Museum

*Laure Philippon*

Institut für Wissensmedien – Knowledge Media Research Center  
Schleichstraße 6, 72076 Tübingen, Germany  
(l.philippon@iwm-kmrc.de)

*Katharina Scheiter*

Institut für Wissensmedien – Knowledge Media Research Center  
Schleichstraße 6, 72076 Tübingen, Germany  
(k.scheiter@iwm-kmrc.de)

*Daniel Bodemer*

University of Duisburg-Essen  
Lotharstr. 65, LE, 47057 Duisburg, Germany  
(daniel.bodemer@uni-due.de)

**Abstract.** The present study is centered on a multi-touch table (MTT), which provides museum visitors with information concerning art exhibitions. The MTT offers the possibility to assist visitors in an either implicit or explicit way to collaboratively engage in meaning-making activities. The present study investigates which kind of support is needed and how it should be designed to foster learning. An explicit instructional support will be designed against the backdrop of research on example comparisons and schema acquisition. An implicit support will be designed as it is known into the field of CSCL. Both ways of support are based on ontological commonalities and differences between exhibits so that the MTT is able to make museum visitors aware of the conceptual structure of the assembly of art exhibits. The two kind of support will be empirically investigated in this (planned) study.

**Keywords:** CSCL, Museum, Instructional Support, Art Exhibitions

## Problem Definition

In the last decade, museums have become an essential context of informal learning for visitors of all ages. In fact museums act as agents of knowledge transfer. However, relatively little is known about how visitors learn in museums and how they can be supported. That is, how can knowledge be efficiently constructed at the museum? More and more museums install multimedia technology in their exhibitions (e.g. the MoMA's iPhone App<sup>7</sup>). Indeed digital technologies may offer new possibilities to learn at the museum. In particular multimedia representations (i.e., combinations of visual and verbal representations in either static or dynamic formats), can be made available to individual visitors as well as to groups visiting a museum. The potential of learning with multimedia is documented within the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2005). However studies investigating the concept of learning with multimedia *at* the museum are still lacking.

The present study is part of an interdisciplinary project aiming at the development of an innovative visitor information system centered on a multi-touch table (MTT), which provides visitors with information concerning art exhibitions. While going through the exhibition, they assemble exhibits that they find interesting with their iPhone by simply selecting them on their display. After their visit, they can survey their collection by depositing the iPhone on the MTT. Their assemblage then appears automatically in a personal area on the MTT display, as shown in Figure 1.

While the original exhibits are still hanging untapped, it is possible to “touch”, to pivot, to displace or to zoom in the collected exhibits on the MTT. Moreover and in contrast to audio guides, this approach allows visitors to see and touch the collection of other visitors, thereby informing them about further art features. The display of divergent ways “to see art” – relative to their own collection – stimulates interaction among visitors and engages them to discuss their different opinions.

---

<sup>7</sup> <http://www.moma.org/explore/multimedia/> [16.12.2011]

Accordingly, this approach accounts better for the fact that visitors rarely go alone to the museum to see an exhibit. Whether school classes, families or guided tour groups, it seems that making the experience of an exhibit *with others* is preferred. In line with this, the present study aims at developing applications for the MTT, which support collaborative learning at the museum.

### **Supporting Collaborative Learning at the Museum**

Collaborative learning has already been well put into practice in school and organizational learning settings but is still lacking in the field of learning at the museum. Anyhow, we think that supporting collaborative learning can be implemented there in fruitful ways, whereby the present application focuses on two approaches: the notion of group awareness and the concept of schema acquisition (comparison of commonalities and differences). Both will be illustrated in the following.

Based on the idea that groups benefit when members are informed about the status of other group members, group awareness tools gather context information about group members (e.g. who is online) and feed this information back. With regard to computer-supported collaborative learning, particularly providing cognitive information, such as knowledge or hypotheses of group members proved to be beneficial (e.g., Bodemer, 2011). For supporting collaborative visitor learning in the museum, gathering and feeding back cognitive variables (such as preferences concerning art exhibits) might be used to initiate meaningful discussions on controversial visitor perspectives. For instance, as shown in Figure 2, both visitors have assembled different exhibits: the MTT displays a luminous border around concordant exhibits to make visitors aware of the conceptual (dis-)similarities regarding their assemblies. In that way, the application implicitly support visitors to discuss their opinion and exchange their knowledge related to art.

Another widely proposed instructional method for fostering learning is the comparison of (multiple) examples (e.g. Quilici & Mayer, 1996; Scheiter, Schuh, & Gerjets, 2004). Comparison processes are seen as essential for schema acquisition (Gerjets et al., 2008). Comparing examples is assumed to allow for the identification and abstraction of relevant features that make up concepts – here concepts of art (e.g. epochs, genre or themes). Such comparison processes may occur spontaneously; however, research suggests that learners typically benefit from instructional guidance that prompts them to conduct comparison processes. According to this, our application consists in an explicit instruction, which prompts the visitors to compare their collected exhibits. For example, the art topic “landscape” can be illustrated both with paintings from different epochs or with paintings of landscapes in different art genres. One visitor may have collected a vase whereas another one has picked a canvas – both displaying a landscape. Thus, they may share a non-obvious interest in landscapes, which they can deepen by comparing and discussing the assembled exhibits.

### **Method**

The aim of this study is to investigate which kind of support is needed for collaborative learning at the museum and how it should be designed for an art exhibition. In the planned study, learners will first work on a test of knowledge about art facts. Then, they will be instructed to visit an experimental exhibition (reproductions of art exhibits) and collect some exhibits with an iPhone. The material (40 exhibits of the experimental exhibition) has been chosen so that it shows different conceptual categories regarding art (e.g. epochs, genre, landscapes, portraits and so on). Subsequently, two participants will meet at the MTT to take part in a collaborative learning situation, in four different experimental conditions. The conditions derive from a 2x2 factorial between-subjects design with the two factors “implicit support” (highlighting divergent exhibit assemblies; given vs. not given) and “explicit instruction” (instruction to compare exhibits; given vs. not given). In the control/ first condition, participants neither receive implicit support nor explicit instruction. In the second condition, they are made aware of the conceptual commonalities and differences in their assemblies by the

luminous border that highlights exhibits in their assembled exhibits; however, they do not receive an explicit instruction to compare exhibits. In the third condition, learners are not implicitly made aware of the conceptual characteristics regarding their assemblies but an explicit instruction about how to deal with the collected exhibits is given (e.g., “Please compare exhibit A from visitor X with B from Y”). In the fourth condition, learners are given both implicit support concerning the identification of conceptual (dis-)similarities among single exhibits as well as an explicit instruction to compare the assembled exhibits. As dependent variables the degree, content, and quality of learner interaction as well as performance in a post-knowledge-test that assesses the ability to identify and describe relevant art features will be assessed. First results of the study will be presented at the conference.

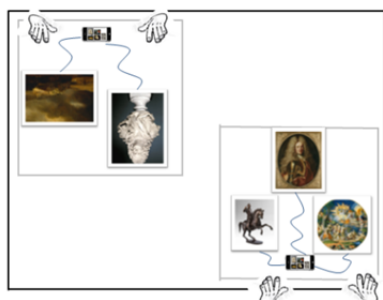


Figure 1. Two visitors with their assembled exhibits in a personal area on the MTT

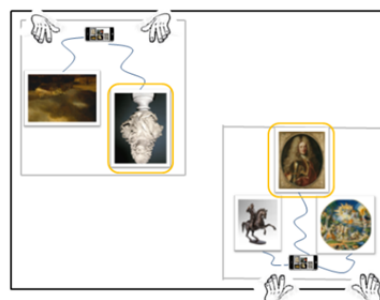


Figure 2. The MTT displays a luminous border around similar exhibits

## References

- Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27, 1097-1086.
- Gerjets, P., Scheiter, K. & Schuh, J. (2008). Information comparisons in example-based hypermedia environments: Supporting learners with processing prompts and an interactive comparison tool. *Educational Technology, Research & Development*, 56, 73-92.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.
- Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problem. *Journal of Educational Psychology*, 88, 144-161.
- Scheiter, K., Gerjets, P., & Schuh, J. (2004). The impact of example comparisons on schema acquisition: Do learners really need multiple examples? In Y. B. Kafai, B. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the 6th International Conference of the Learning Sciences* (pp. 457-464). Mahwah, NJ: Erlbaum.



# Fourth Graders' Text and Picture Integration in processing and Learning from Science Text: Evidence from Eye Movement Patterns

*Patrik Pluchino, Maria Caterina Tornatora, Lucia Mason*

University of Padova, Italy  
Via Venezia 8, 35131 Padova, Italy  
patrik.pluchino@unipd.it

**Abstract.** This study used eye-tracking methodology in the school setting to examine fourth graders' online processing of an illustrated science text. We were interested in identifying patterns of visual behaviour, which was examined considering some individual differences. We also investigated learning from text as well as the link between processing and learning. Forty-nine 4<sup>th</sup> graders participated in a pretest, immediate, and delayed posttest design. Cluster analysis revealed three patterns of eye movements varying for the level of integration of text and picture. Significant associations between eye-fixation data and reading comprehension and prior knowledge emerged. Moreover, the patterns of visual behaviour were significantly related to learning outcomes. Students who performed better in the various tasks were those who showed greater integrative processing of the illustrated text.

**Keywords:** Text processing, eye movements, text and picture comprehension, learning from text and picture

## Introduction

Successful comprehension of the learning material that is usually presented in textbooks requires comprehension of verbal and graphical information, especially in science education. Research has investigated the potential benefits of multiple external representations when learning complex concepts (Ainsworth, 1999, 2006). It has been documented that pictures enhance students' learning from text (Carney & Levin, 2002). For science learning, especially at elementary and middle school levels, visualizations play a crucial role as they can make unseen and complex processes visible (Kali & Linn, 2008). However, research has also indicated that it is not always true that "a picture is worth a thousand words" (Mayer & Sims, 1994). What does comprehension of an illustrated text require? According to the Schnotz (2002) integrative model, which elaborated Mayer's (2001) theory of multimedia learning, structural mapping processes are essential to the formation of a coherent mental model of an illustrated text. Such a model implies the integration of the information provided by the text with information provided by the picture (Ainsworth, 2006; Florax & Ploetzner, 2010; Schlag & Ploetzner, 2011; Schnotz & Bannert, 2003).

Eye-tracking methodology is useful to shed light on the process of illustrated text comprehension. Some eye-tracking studies have contributed to unveiling text and picture processing in college students (e.g. Hegarty & Just, 1993; Sanchez & Wiley, 2006), but to our knowledge only a very few studies have involved younger learners (Hannus & Hyönä, 1999; Mason, Pluchino, Tornatora, & Ariasi, in press). In this investigation we were particularly interested in tracking the integration of verbal and pictorial information through eye movement analyses in elementary school students, that is, in much younger learners than those involved in previous research. We also sought to investigate the link between the online processing and the offline outcomes of learning from it. More specifically, to extend previous investigation, the following research questions guided the present study: (1) Are different patterns of readers' eye movements during the first and second-pass reading identifiable? (2) Are eye movements related to the individual characteristics of reading comprehension ability, prior knowledge, and spatial ability? (3) Is readers' learning from and illustrated science text associated to their eye movement patterns during reading, after controlling for individual differences?

## Method

### *Participants.*

Forty-nine 4<sup>th</sup> grader (F=22, mean age 9,7 years) were involved in a pretest, posttest, and delayed posttest design.

### *Learning material.*

All participants read the same text on the topic of air. The text presents the characteristics of air and then explains that there is air even when it seems that there is nothing. The text then describes a phenomenon to explain why an empty water bottle contains air that must escape in order for water to be poured into it. An illustration depicts the phenomenon. The text and illustration were presented on the same computer screen, taking into account the spatial contiguity principle (Ginns, 2006).

*Reading comprehension* was measured using the MT (Italian) test for fourth grade (Cornoldi & Colpo, 1998).

*Prior knowledge* about the topic of air at pretest was assessed by five open-ended questions. Inter-rater reliability for each question ranged from .90 to 1 (Cohen's  $k$ ). *Spatial ability* was measured using the subtest for spatial relations of the Primary Mental Ability Test (PMA, Thurstone & Thurstone, 1963).

### *Learning tasks.*

Free recall, factual knowledge, and knowledge transfer were measured at both immediate and delayed posttests. *Eye-Fixation Measures.* Five indices were computed considering the various AOIs (Hyönä & Lorch, 2004; Hyönä, Lorch & Rinck, 2003):

1. First-pass fixation time on text (duration of all fixations during the first-pass reading);
2. first-pass fixation time on illustration (duration of all fixations during the first-pass inspection);
3. text look-from fixation time (duration of all re-fixations that “took off” from a text segment and “landed” on another text segment);
4. look-from text to illustration fixation time (durations of all re-fixations that “took off” from a text segment and “landed” on the illustration);
5. look-from illustration to text fixation time (duration of all re-inspections that “took off” from the illustration and “landed” on a text segment). All these measures were logarithmically transformed due to the large variance in participants' visual behaviour that led to non-normal distributions.

## **Results**

### *Eye Movement Patterns*

A cluster analysis using the Ward method revealed three patterns of processing. The least frequent, pattern 2 ( $n=7$ ; 14%), was characterized essentially by the least time dedicated to the graphical representation and by the least integration of verbal and visual representation during the second pass text reading or picture inspection. The more frequent pattern 1 ( $n=20$ ; 41%) was characterized by less time spent inspecting the picture during the first pass and less time spent integrating verbal and visual information during the second pass. The most frequent, pattern 3 ( $n=22$ ; 45%), was characterized by the longest fixation time on inspecting the picture during the first pass and more integrating text and picture during the second pass. Kruskal-Wallis non parametric tests revealed significant differences between the three patterns in relation to the first-pass fixation time on the picture,  $X^2(2) = 16.32$ ,  $p < .001$ , the look-from text to text fixation time,  $X^2(2) = 16.97$ ,  $p < .001$ , the look-from text to picture fixation time,  $X^2(2) = 26.23$ ,  $p < .001$ , and the look-from picture to text fixation time,  $X^2(2) = 25.99$ ,  $p < .001$ .

### *Eye Movements and Individual Differences*

Correlational analyses revealed significant associations between some of eye-movement data, reading comprehension ability and prior knowledge. Specifically, the lower the general reading comprehension ability, the longer the first-pass fixation time on the text ( $r = -.34$ ). More or less time is required to

attend to text segments by more or less skilled comprehenders, especially when the learning material becomes complex. In addition, it emerged that the higher the level of learners' preexisting knowledge of the topic, the longer their fixation time on the picture during rereading of the text ( $r = .31$ ) and the longer their fixation time on the text during reinspection of the picture ( $r = .30$ ). Higher prior knowledge was therefore associated with a strategic behaviour which is crucial to learning from illustrated text. Spatial ability was not associated with any eye movement measure.

### *Eye Movement Patterns and Learning*

First regression analyses were performed to obtain the residual standardized scores for all the learning measures at both testing times, after controlling for the possible contribution of reading comprehension and prior knowledge to these measures. Successively, we performed Kruskal-Wallis tests with the adjusted standardized scores to evaluate whether the patterns of eye movements associated to the participants' performances in the learning tasks at immediate and delayed posttests. These tests revealed that the patterns of visual behaviour were significantly related to immediate recall,  $X^2(2) = 12.17$ ,  $p = .002$ , immediate factual knowledge,  $X^2(2) = 10.26$ ,  $p = .006$ , delayed factual knowledge,  $X^2(2) = 14.85$ ,  $p = .001$ , and immediate transfer,  $X^2(2) = 6.05$ ,  $p = .048$ . Pattern 3 students obtained higher scores in all these learning tasks, differentiating from pattern 2 and pattern 1 students.

### **Discussion**

As expected, it was the integration between the two types of external representation that differentiates, essentially, the three patterns of eye movements while processing the illustrated text. In agreement with the theoretical models and empirical studies on multimedia learning, pattern 2 appeared to be less supportive and functional to knowledge acquisition from the illustrated text. In contrast, pattern 3 appeared to be the most strategic and functional resource for learning due to greater integration of verbal and graphical information. Pattern 1 appeared to be an intermediate way of dealing with text and picture processing during the first and second pass. Data regarding learning from the illustrated text confirmed that students who spent longer time trying to integrate text and picture outperformed the other students.

The study contributes to the literature on processing and learning from illustrated science texts at primary school level by focusing on multiple indices of eye fixations during first- and second-pass reading or inspection, and taking into account some individual differences. Investigations focused on online cognitive processing for comprehension of text and picture are important to further our understanding, at least to some extent, of the reasons underlying low achievement in science, which is based on successful illustrated text comprehension from the last years of primary school.

### **References**

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16, 183-198. doi:10.1016/j.learninstruc.2006.03.001
- Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313-334). Amsterdam: Elsevier.
- Hyönä, J., & Nurminen, A.-M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British journal of Educational Psychology*, 97, 31-50. doi:10.1348/000712605x53678
- Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (in press). An eye-tracking study of learning from science text with concrete and abstract illustrations. *Journal of Experimental Education*.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Schlag, S., Ploetzner, R. (2011). Supporting learning from illustrated texts: conceptualizing and evaluating a learning strategy. *Instructional Science*, 39, 921-937. doi:10.1007/s11251-010-9160-3
- Schnotz, W. (2002). Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14, 101-120. doi:10.1023/A:1013136727916

# Effects of Simplified Ancillary Representations and Cues on Learning from Animation

Dian-Kemala Putri<sup>1</sup>, Jean-Michel Boucheix<sup>2</sup>, Richard. K. Lowe<sup>3</sup>

<sup>1</sup>University of Gunadarma, Margonda Raya 100 Depok, Indonesia, dian@staff.gunadarma.ac.id

<sup>2</sup>LEAD-CNRS, University of Burgundy, Dijon, France, Jean-Michel.Boucheix@u-bourgogne.fr

<sup>3</sup>Curtin University, Perth, WA, Australia, R.K.Lowe@curtin.edu.au

**Abstract.** Effects of novel presentation formats on learning from animation were tested in 2 experiments. Experiment 1 recruited Johnson Laird's notions about simplification in mental model building and compared sequential combinations of different formats (static and animated) versus repeated presentation of single formats. Comprehension measures indicated that despite participant's mental model quality being significantly higher for the static-animated sequence, mental model scores were relatively low. Experiment 2 aimed to raise mental model quality by adding a new form of entity cueing in which functionally important aspects were cued via coloured tokens. Although eye movement data indicated the cues were effective in directing learners' attention they did not result in further improvements in mental model quality.

**Keywords:** animation, mental model building, simplified representation, token cueing, eye tracking

## Introduction

Building high quality mental models from an external animation can be a demanding activity. Johnson-Laird (1998) argued that to be effective, mental models need to be "*simplified representations*" or "*small-scale models*" of the external world. Accordingly, some of the difficulties learners have in building effective mental models from animations may arise because they fail to simplify or scale down the referent content before attempting to internalize it. The two experiments reported here examined whether the chaining of simplified external representations to animated presentations and the application of cues fostered building of more effective mental models. Simplified formats such as static pictures or animation segments depicting key steps of an animation have been used in previous studies (Kriz & Hegarty, 2007; Arguel & Jamet, 2008; Bétrancourt & Morand, 2010). However, when such ancillary representations were presented in isolation or simultaneously with the animation, the results were inconclusive. It may be that these approaches lack sufficient links between the simplifications and the animations or impose excessive processing demands on the learner. We took an alternative two-stage approach to presenting simplified external representations which manipulated the presence and the temporal location of (i) a set of static pictures and (ii) the corresponding animation. A complex animation of an upright piano mechanism was used without text (Figure 1). Four pairings of presentations were compared. In the first condition (**SA**), a set of six Static pictures (**S**) depicting key steps of the mechanism was presented and then followed by the Animation (**A**). In the second condition, the presentation order was reversed (**AS**). The two other conditions were single format controls (**SS** and **AA**). Higher quality mental models were predicted in the **SA**, and to a lesser extent in the **AS** conditions, than in the single format conditions. In the **SA** condition, we expected that in the first stage, learners would build a mental model of the mechanism's operation by making inferences between the static key steps via active comparison and mental manipulation of pictures in the set. In the second stage, learners would use the animation as feedback for checking and refining this mental model. In the **AS** condition, learners could apply the mental model built from the animation to help them infer the changes depicted in the static pictures. Although this would allow them to try-out the mental model they constructed from the animation, the static pictures would be less useful for checking and refinement. In the single format conditions, there are no such opportunities to test the first-stage mental model against an alternative representation.

## Experiment 1

After testing for spatial abilities and prior knowledge, the 82 participants (undergraduate French students) were randomly assigned to four groups (two dual-format and two single-format conditions). Each participant studied how the piano mechanism worked from the depictions provided (Fig. 1) for a total learning time of 3'30", (1'45" for each stage). A local motion post-test measured recall of the local motion

of each piano component. In this test, learners used the computer mouse to move crosses on a static picture of the mechanism to show the direction and amplitude of each component's movement. A second post-test, in which participants were asked to write an accurate explanation of the piano mechanism's functioning, measured the level of integration of the "conceptual model" they had developed. Comprehension scores are presented Table 1a.

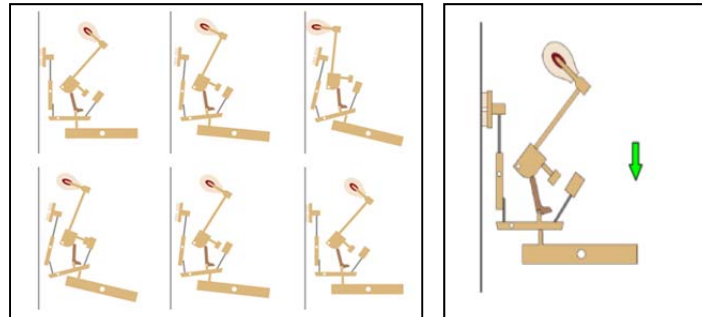


Figure 1. Series of static pictures and animation (SA) (Experiment One)

A repeated measures ANOVA showed mainly a significant interaction between type of presentation and post-test,  $F(3, 78) = 6.52, p < .001, \eta_p^2 = .20$  favouring the **SA** and the **AS** conditions for the conceptual model measure only. Univariate analysis showed an effect of presentation type for the conceptual model score,  $F(3, 74) = 4.44, p = .006, \eta_p^2 = .15$ , with scores for **SA** and **AS** being higher than those for **SS** and **AA**,  $F(1, 74) = 10.01, p = .002$ . However, there was no effect of presentation type for local motion scores,  $F(3, 74) = 0.85, ns$ . Finally, local motion scores were higher than conceptual model scores  $F(1, 74) = 290.05, p < .001, \eta_p^2 = .79$ .

## Experiment 2

Experiment 1 showed participant's integrated conceptual model was significantly higher for the dual format presentations than for the single format presentations. However, even in the best condition (**SA**), conceptual models scores were relatively low (Table 1). In experiment 2, entity-based cueing was used with the aim of directing learner's attention to specific areas with particular relevance for building high quality conceptual models. A novel form of dynamic cueing (see Boucheix & Lowe, 2010 and also the review by De Koning & al., 2009) was adapted for simplified representations by localizing cueing in functional groups of mechanism entities using coloured *tokens* (Fig. 2). The use of such tokens across key steps of the dynamic process could further improve the quality of conceptual models resulting from the chaining of simplified representations and animations.

Participants were 89 French undergraduate students. The experimental conditions and post-tests were the same as in experiment 1, except for the use of token cues, addition of a control condition without cues, and inclusion of eye tracking measures (Tobii 120). Based on the experiment one results, we choose the **SA** condition as the un-cued control condition for experiment 2. The same set of six static pictures was replicated four times, with each replica cueing a different functional group in its own colour. These displays were presented for 26 seconds per screen. The cue presentation order was consistent with the progressive changing dynamics of the piano operation (1: green, key-whippen; 2: red, jack-hammer butt; 3: blue, damper-hammer and 4: yellow, balance-back check.).

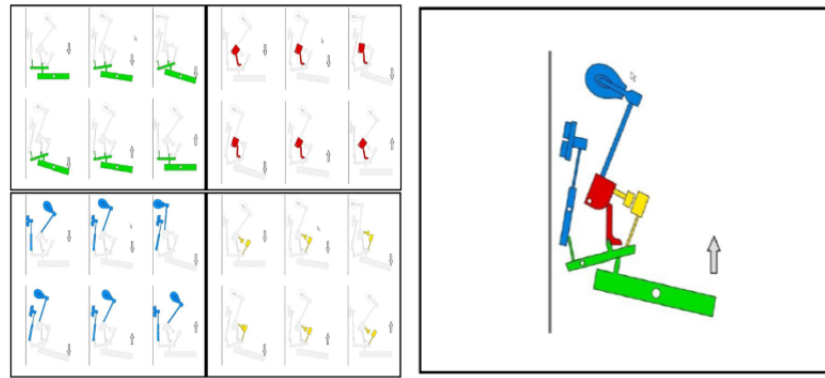


Figure 2. An example of **SA** condition (Experiment 2)

Table 1a, and 1b- Mean score as percentage of correct answers (and SD), for local motion and mental model quality tests, for each type of presentation, table 3a experiment1, and table 3b experiment 2.

1a, Exp. 1	SA	AS	SS	AA	1b, Exp. 2	SA	AS	SS	AA	Control
Local	41.8	38.8	42.2	45.6	Local	46.2	47.7	45.8	49.3	50.3
Motion	(14.9)	(13.9)	(8.01)	(13.4)	Motion	(7.6)	(9.9)	(12.0)	(9.9)	(13.8)
Conceptual	24.8	20.8	16	13.3	Conceptual	38.9	40.6	35.2	43	35.3
Model	(12.8)	(15.7)	(9.41)	(9.7)	Model	(13.7)	(16.6)	(17.9)	(15.2)	(12.5)

Comprehension scores are presented table 1b. A repeated measures ANOVA showed no significant effect of presentation type,  $F(4, 79) = 1.2, p = 0.3, \eta_p^2 = .05$ , and no significant interaction. There was no significant difference between the groups in the cued **SA** and un-cued control conditions for the both post-tests,  $F(1, 79) < 1, ns$ . Analysis of eye movement data showed that participants looked significantly more often and longer at areas with tokens than at the rest of the display. Fixations on these areas persisted from when the token cue appeared until its disappearance indicating that the cueing directed the learner's attention efficiently (despite not resulting in comprehension benefits).

## Conclusion

The addition of a simplified external representation chained to an animation resulted in superior conceptual model quality. No further benefit was obtained by adding entity-based token cueing. A possible explanation for this lack of benefit is that although these cues are effective in directing attention to an animation's functionally relevant entities, they fail to signal the key events that relate these entities and are crucial for the building mental models. However, the absolute values for conceptual model seem different between experiments 1 and 2, although the **SA** condition in exp. 1 and the **SA** control in exp. 2 were the same. Since cueing had no effect, one would expect to replicate findings from Exp. 1 in Exp. 2. Further research is needed to better understand such inconsistency.

## References

- Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior*, 25(2), 354-359.
- Boucheix, J.M. & Lowe, R.K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning from complex animations. *Learning and Instruction*, 20, 123-135.
- De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21, 113-140.
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, 65, 911-930.
- Johnson-Laird, P. N. (1998). Imagery, visualization, and thinking. In J. Hochberg (Ed.), *Perception and Cognition at the Century's End* (pp. 441-467). San Diego, CA: Academic Press.
- Morand, L. & Bétrancourt, M. (2010). *Collaborative Learning with Single or Multiple Animations. EARLI Conference, SIG2. Comprehension of Text and Graphics, Tübingen, August 25-28.*

# Visual Ability in Navigation Communication

*Annemarie Quispel<sup>1</sup>*

*Alfons Maes*

<sup>1</sup>Avans University, Tilburg Center for Cognition and Communication (TiCC), The Netherlands  
a.quispel@uvt.nl

**Abstract.** Recent studies have shown that there are two distinct types of visual intelligence. Visual-object intelligence reflects one's ability to process information about visual appearances of objects, whereas visual-spatial intelligence reflects one's ability to process spatial relations and manipulate objects in space. These studies also showed that visual-object intelligence is related to specialization in visual art and design, whereas visual-spatial intelligence is related to specialization in engineering/sciences. The aim of this study is to show that the two types of visual ability result in different strategies for processing and producing route information.

**Keywords:** visual intelligence, visual ability, object visualizer, spatial visualizer, route information, navigation

## Introduction

Several studies in the past decade provided support for a distinction between two types of individuals based on their imagery capacities. Kozhevnikov, Kosslyn, Shephard (2005) identified individuals with high object-imagery ability, called object visualizers, and individuals with high spatial-imagery ability, called spatial visualizers. Results indicated that object visualizers encode and process images holistically, as single perceptual units, whereas spatial visualizers generate and process images analytically, piece by piece. Blazhenkova & Kozhevnikov (2010) showed that visual-object intelligence is a unique component of intelligence, which reflects one's ability to process information about visual appearances of objects and their pictorial properties (e.g., shape, color and texture), and demonstrated that this is distinct from visual-spatial intelligence, which reflects one's ability to process information about spatial relations and manipulating objects in space (e.g. mental rotations).

According to both Kozhevnikov et al. (2005) and Blazhenkova et al. (2010) visual-object ability relates to specialization in visual art, whereas visual-spatial ability is related to specialization in science (physics, biochemistry, engineering, computer science and mathematics).

## Research questions

The main research questions were (1) whether the distinction between visual-object and visual-spatial ability is indeed reflected in the distinction between visual artists/designers and engineers/scientists, and (2) how these types show up in multimodal navigation tasks. In particular, we wanted to find differential traces of spatial and object abilities in the way these learners produce route descriptions and route drawings, thereby controlling factors we expect to be closely related to these ability types, like perspective (route vs. survey) or route shape.

Route information can be described from either a survey perspective (with an extrinsic frame of reference, map-like) or a route perspective (with an egocentric frame of reference, like ground-level navigation). When producing route descriptions based on route maps, we expect object visualizers to take a survey perspective and to exploit route shapes more often, and to focus more on discrete entities on the map (e.g. landmarks along the route), whereas spatial visualizers are expected to exploit a route perspective more often and to focus more on path entities (e.g. streets, patterns). When asked to draw routes on a predefined map, we expect spatial visualizers to outperform object visualizers when drawings are based on route (as opposed to survey) descriptions, as route perspective involves operations of mental rotation. Furthermore, when asked to draw a map based on a description including path entities and objects, we expect object visualizers to focus more on object details and spatial visualizers more on path details. Additionally, we wanted to explore other differences between the learners in drawing based on route or survey descriptions (e.g. level of detail and drawing time).

## Method

### *Participants*

Sixteen bachelor students in art and design (9 male, 7 female, mean age 23) and sixteen bachelor students in engineering (15 male, 1 female, mean age 21) of Avans University of Applied Sciences volunteered to participate in the experiment. Participation was individual.

### *Design and materials*

The experiment consisted of three tasks (1. controlled route production, 2. controlled route drawing, 3. 'free' map drawing). Learner type (designers vs. engineers) was a between participants factor, perspective (survey or route) was a between participants factor in task 2 and 3, divided evenly over the learner types; route shape (regular or irregular) was a within participants factor in task 1.

For the controlled route production and drawing task, a set of 16 routes was designed, varying in shape (8 routes taking the shape of an object, like a square box or stairs, 8 irregular shaped), length (4 or 6 segments), route orientation (left-right, top-bottom and vice versa) and number of turning points (2 to 5). Each route was drawn on an identical grid of squares, each turning point being flagged by external landmarks in the form of written labels (e.g. "cafe", "gas station"). For all route maps two route descriptions were constructed for the controlled drawing task, one in survey, one in route perspective. The descriptions were recorded in two sets of 16 spoken directions: one from a survey and one from a route perspective. (See Fig.1 for examples.) For task 3, two descriptions were composed of a small town, one from a survey and one from a route perspective, of comparable length (137 and 132 words).

### *Procedure*

Participants took the tests sitting in a quiet room behind a computer, pressing the space bar to proceed through the instructions in task 1, and with pen and paper for completing task 2 and 3.

*Task 1.* Each participant was presented with the 16 route images, one by one in random order, and asked to describe each route such that someone else could easily choose it from a set of route maps.

*Task 2.* Each participant was presented with spoken descriptions of the 16 routes, one by one in random order, either from a survey or from a route perspective. They were asked to draw each described route (pen on paper) on a predefined map with only a starting point depicted on it.

*Task 3.* Each participant was presented with a description on the computer screen of a small town, either from a route or from a survey perspective (for each participant, the perspective in task 2 and 3 alternated). They were asked to draw a map of the described town, taking as much time and looking back on the screen as often as they needed.

*Visual ability tests.* Subjects' visual ability was tested using the Object-Spatial Imagery Questionnaire (OSIQ: Blajenkova et al., 2006; Dutch translation by Diane Pecher), the Degraded Pictures Test (Ekstrom, French, Harman, 1976), and the Spatial Imagery Ability Test (MMVirtualDesign).

## Results

### *Visual ability tests*

Designers scored higher on the OSIQ object scale than engineers ( $t(30)=3.05$ ,  $p<.005$ ), while engineers scored higher on the spatial scale than designers ( $t(20) = 4.23$ ,  $p < .001$ ). The designers scored higher than the engineers on the Degraded Pictures Test ( $t(30) = 2.03$ ,  $p < .05$ ). Unexpectedly, the Spatial Imagery Ability Test showed no difference between the two groups, except in executing time: designers took much more time to complete it than engineers ( $t(22) = 2.05$ ,  $p < .05$ ).



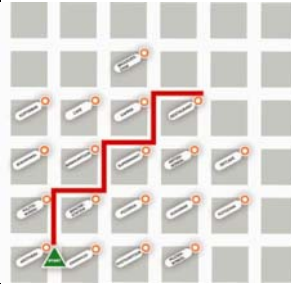
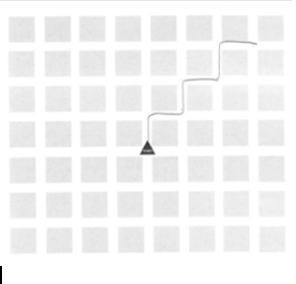
Task 1	Task 2		
Production of descriptions	Drawing task - Survey version	Drawing task - Route version	Drawing result
	Go one block up. Go one block to the right. Go one block up. Go one block to the right. Go one block up. Go one block to the right.	Go right at the next intersection. Go left at the next intersection. Go right at the next intersection. Go left at the next intersection. Go right at the next intersection. Go to the next intersection.	

Figure 1. Task 1: one of 16 maps (6 segments, 5 turning points, shape); Task 2: English translation of spoken directions for this map (survey and route version) and an example of a (correct) drawing.

*Task 1: descriptions.* Almost all subjects ( $N = 29$ ) showed a preference for using one perspective consistently. Nine subjects consistently used the survey perspective, using expressions like ‘down’ and ‘up’, and ‘to the left/right’ indicating literally to the left or right of the plane. More designers than engineers (7 vs. 2) used the survey perspective. As expected, the mean number of routes including discrete entities (landmarks) was higher for the designers than for the engineers ( $t(23) = 1.78, p < .05$ ). On the other hand, the mean number of routes including path entities (references to streets) was higher for the engineers than for the designers ( $t(17) = 1.96, p < .05$ ). Respondents hardly used any reference to the shapes of the routes.

*Task 2: route drawing.* More errors were made in drawings based on route (as opposed to survey) perspective ( $t(30) = 4.62, p < .001$ ). Scores on the OSIQ spatial scale correlated with scores on drawing correctness ( $r = .42, p < .05$ ). There was no difference between the two groups, though.

*Task 3: map drawing.* The designers drew more detailed, iconic landmarks than the engineers: 27 vs. 14 (designers:  $M = 1.63, SD = 2.6$ ; engineers:  $M = .88, SD = 1.7$ ). The designers took more time to draw than the engineers (designers:  $M = 168, SD = 72$ ; engineers:  $M = 133, SD = 48$ ).

## Discussion

The results from the visual ability tests, specifically the OSIQ, confirm Blahzenkova and Kozhevnikov’s findings that object and spatial imagery are related to specialization in visual arts and sciences respectively. The equal scores between the two groups in the Spatial Imagery Ability Test calls for further analysis of the relation between professions and spatial ability tests. Altogether, the findings show that differences in visual ability lead to differences in processing and producing route information. Designers, associated with object imagery, tend to process this from a survey perspective and focus more on objects (landmarks), while engineers, associated with spatial imagery, tend to process this more from a route perspective, step-by-step, focusing on paths (street patterns).

## References

- Blahzenkova, O., Kozhevnikov, M. (2010). Visual-object ability: A new dimension of non-verbal intelligence. *Cognition* 117 (pp. 276-301).
- Kozhevnikov, M., Kosslyn, S., Shepard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition*, 33 (4) (pp. 710-726).

# Teenage Students' Awareness of Source Quality when Selecting Web Links

Jean-François Rouet<sup>1</sup>, Christine Ros<sup>1</sup>, Guillaume de Pereyra<sup>1</sup>, Mónica Macedo-Rouet<sup>2</sup>, Ladislao Salmerón<sup>3</sup>

<sup>1</sup>CNRS and University of Poitiers, France, <sup>2</sup>CNDP and University of Poitiers, France

<sup>3</sup>University of Valencia, Spain

jean-francois.rouet@univ-poitiers.fr

**Abstract.** This exploratory study examined the criteria used by teenage students' when selecting documents in a simulated Web search environment. Forty-eight 7th graders and 47 9th graders were given simple search topics and asked to select the most useful item in a list of four simplified Web links made of content and source information. Topical relevance and source authority of the links were manipulated in a within-subject design. The participants also took a reading fluency test and an Internet-specific epistemic beliefs inventory. Ninth graders were better at taking into account the authority of the source in their selections. They were also more likely to cite source features when justifying their selections. Neither reading fluency nor epistemic beliefs were found to significantly relate to performance. Implications for future research are briefly discussed.

**Keywords:** teenagers, source, relevance, credibility, Web search

## Introduction

Internet-based inquiry tasks have become increasingly popular in primary and secondary education. However, acquiring quality information through the Web is a notoriously difficult task for most students. The present study explored the criteria used by 7th and 9th graders when selecting links in a simulated Web search environment. The study also examined the consistency of students' epistemic beliefs and the relationship between reading fluency, epistemic beliefs and link selection strategies.

### *Students' assessment of topical relevance and source credibility*

When selecting links in a search engine results page, Web users may consider both the content of the link (does the link match the search topic?) and its source (does the information come from a reliable source). The evidence suggests that the acquisition of quality information through the Web is a challenging task for teenage students. Students' judgements are often based on superficial or irrelevant cues. For instance, Eastin, Yang, and Nathanson (2006) found that 3rd, 4th and 5th graders found Internet websites more credible if the site contained dynamic graphics than if they did not. Rouet, Ros, Goumi, Macedo-Rouet, and Dinet (2011) observed that 5th to 9th graders tend to select titles based on keywords or typographical emphasis, as opposed to the meaning of the title.

Research on grade school students' assessment of the credibility of information sources is still scarce. Macedo-Rouet, Braasch, Britt, & Rouet (subm.) found that 4th and 5th graders seldom mentioned sources features even when asked to make a judgment about the knowledgeability of the source. Students' awareness of source features such as authors' credentials seems to emerge toward the end of secondary education (Brem, Russell, & Weems, 2001; Britt & Aglinskias, 2002; Kammerer & Gerjets, 2012). However, there has been thus far no systematic investigation of teenage students' ability to consider source authority in addition to topical relevance as part of a Web link selection task.

### *Rationale of the experiment*

The main purpose of our experiment was to explore 7th and 9th grade students' criteria when selecting links in a simulated Web search environment. Based on available evidence, we anticipated a shift in older students' strategies toward a better consideration of source features, perhaps in relation with students' emerging conceptions of the Internet as an information resource (Eastin, 2008; Strømsø & Bråten, 2010). We used a subscale of Strømsø and Bråten's (2010) ISEQ inventory, as adapted by

Kammerer and Gerjets (2012), in order to explore the consistency of teenage students' responses and their possible link with their selection criteria. In addition, we measured the participants' reading fluency in order to assess the likely contribution of this factor to their performance in the selection task.

## **Method**

### *Participants*

Participants were 48 7th grade students (age range: 12-13 years) and 47 9th grade students (age range: 14-15 years) from a French suburban middle school. The students participated as part of regular computer class sessions.

### *Materials*

Two practice and 10 critical search units were designed. Each search unit was made of one search topic (e.g., "*The history of the China wall*") and 4 links representing potential documents. Each link contained a title and a source description. The titles were either relevant (e.g., "*The evolution of the China wall from its origins*") or irrelevant to the search topic (e.g., "*The history of porcelain: From Limoges to China*"). The source descriptions were either authoritative (e.g., "*historical and tourist guide by the China Information agency*") or less authoritative (e.g., "*personal blog by the owner of a camping ground*"). Titles and source descriptions were assembled so that only one of the four links was both relevant and authoritative. In addition, two filler units were created in which none of the links had a relevant content. Most of the materials were adapted from our prior work with younger students and had been rated as familiar by teachers (see Rouet et al., 2011). Thus, the materials were not intended to present any challenge from the point of view of prior knowledge or vocabulary.

For each unit, the search topic and four links were presented on a computer screen using a simulated search engine page. The students were asked to write down on a printed form the number of the most useful link, or to select "none" if they thought that none of the links was useful. On two trials, they were also asked to provide a short justification for their selection. The order of presentation of units and links within units were counterbalanced across students.

In addition to the search materials, we used a standard French fluency test (Lefavrais, 1986) and the adapted ISEQ certainty subscale<sup>8</sup>.

### *Procedure*

The students participated in two sessions. In the first, 30-minute session, they took the reading fluency test and the ISEQ subscale. In the second, 50-minute session, they took the link selection task and two follow-up tasks that are not reported in this presentation. There was not strict time limit imposed on the selection of links. The main dependent variable was the number of trials in which students selected the relevant and authoritative, or "optimal" link.

## **Results**

Ninth grade students selected an average of 6.81 optimal links (sd=1.86) while 7th graders selected only 5.87 (sd=2.07). The difference was statistically significant ( $t(93)=2.33, p<.05$ ).

In addition, we analyzed a sample of 70 items randomly drawn from the pool of justifications provided by the participants. Ninth graders were more likely to refer to source features as opposed to content features or other types of justifications compared to 7th graders (55% vs. 9%, respectively,  $X^2(3)=14.1, p<.001$ ).

---

<sup>8</sup> we thank Yvonne Kammerer for communicating her version of the epistemic beliefs subscale.

The internal consistency of the ISEQ subscale as assessed by Cronbach's Alpha was .70 at the 7th Grade and .76 at the 9th Grade. These figures are comparable to that reported by Kammerer and Gerjets (2012) and suggest that the participants were not responding randomly to the items.

Correlational analysis between the reading fluency task, the ISEQ subscale and the selection score revealed positive but low and nonsignificant correlations ( $0 < r < .20$ ). Thus, controlling for grade we did not get evidence that students' performance at the link selection task could be explained by either their reading fluency or by their beliefs regarding the certainty of Internet information.

## Discussion

As expected, 9th graders were more likely than 7th graders to take into account source reliability in their selection of a link. They were also more likely to refer to course characteristics when explaining their selection of a link. This pattern of results is consistent with Macedo-Rouet et al.'s (subm.) observation that primary school children seldom refer to source information when justifying source knowledge assessments.

The lack of a significant correlation between Internet-related beliefs and students' selection performance contrasts with prior findings by Strømsø and Bråten (2010) and Kammerer and Gerjets (2012). However these experiments involved older participants and more complex Internet reading tasks. Thus, besides the general observation that students' proficiency improves with schooling, the cognitive and language factors that underlie performance remain to be elicited. Future experiments will assess students' explicit knowledge and evaluation of source features as a possible mediating factor.

## References

- Brem, S.K., Russell, J., & Weems, L. (2001). Science on the Web: Students' evaluation of scientific arguments. *Discourse Processes*, 32, 191-213.
- Britt, M.A., & Aglinskas, C. (2002). Improving students' ability to identify and use source information. *Cognition and Instruction*, 20, 485-522.
- Eastin, M.S. (2008). Toward a Cognitive Development Approach to Youth Perceptions of Credibility. In M.J. Metzger and A.J. Flanagin (Eds.) *Digital Media, Youth, and Credibility* (pp. 29–48). The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press. doi: 10.1162/dmal.9780262562324.029
- Eastin, M.S., Yang, M.-S., & Nathanson, A.I. (2006). Children of the Net: An Empirical Exploration Into the Evaluation of Internet Content. *Journal of Broadcasting & Electronic Media*, 50, 211-230.
- Kammerer, Y., & Gerjets, P. (2012). Effects of search interface and Internet-specific epistemic beliefs on source evaluations during Web search for medical information: An eye-tracking study. *Behaviour & Information Technology*, 31, 83-97.
- Lefavrais, P. (1986). *La pipe et le rat [Pipe and rat]*. Issy-les-Moulineaux, France: Etablissements d'Applications Psychotechniques..
- Macedo-rouet, M., Braasch, J.L.G., Britt, M.A., & Rouet, J.F. (subm.). Teaching fourth and fifth graders to evaluate information sources during text comprehension.
- Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, A., & Dinet, J. (2011). The influence of surface and deep cues on grade school students' assessment of relevance in Web menus. *Learning and Instruction*, 21, 205-219.
- Strømsø, H.I., & Bråten, I. (2010). The role of personal epistemology in the selfregulation of internet-based learning. *Metacognition and Learning*, 5, 91-111.

# Selecting Pages from Google to Learn about a Controversial Topic: The Role of Epistemic Beliefs

*Ladislao Salmerón*

ERI Lectura, Universitat de València  
Avda. Blasco Ibáñez, 21. 46010 Valencia, Spain  
ladislao.salmeron@uv.es

*Yvonne Kammerer*

Knowledge Media Research Center  
Schleichstrasse 6. 72076 Tuebingen, Germany  
y.kammerer@iwm-kmrc.de

**Abstract.** The Web usually provides information of highly variable quality when it comes to controversial topics. This is particularly problematic because many students access mostly pages displayed on top of the search engine results list (SERP) without further consideration about their quality ('Google trust' heuristic). We studied how epistemic beliefs (EBs), that is, personal beliefs about the nature of knowledge and knowing, shape students selection of web pages about a controversial topic. Results showed that students who believed that knowledge is constructed by the self rather than coming from experts visited more often the pages on top of the SERP, and bookmarked pages from this range. On the contrary, students that believed that knowledge claims need to be justified through rules of inquiry and cross-checking of different sources accessed more pages from the middle and bottom of the SERP, and weren't influenced by the 'Google heuristic' to bookmark their pages.

**Keywords:** reading comprehension, Web navigation, 'Google trust' heuristic, epistemic beliefs

## Introduction

Students in higher education often use the Web to learn about controversial topics, such as climate change. One of the main steps involved in learning from the Web involves the selection of relevant pages for further study. Search engines assist students in this process. However, as controversial topics do not convey a single clear-cut solution and as the Web usually provides information of highly variable quality, students should evaluate, compare, and integrate information from multiple websites instead of blindly relying on the first few results provided by the search engine ('Google-trust' heuristic). There is mounting evidence that epistemic beliefs (EBs), that is, personal beliefs about the nature of knowledge and knowing, play an important role in the processing of controversial information from multiple documents (for a review see Bråten, Britt, Strømsø, & Rouet, 2011).

But what are the mechanisms by which epistemic beliefs influence learning and comprehension? Muis (2007) has argued that readers transform their epistemic beliefs into 'specific epistemological standards' that they use to qualify their self-regulated learning processes such as task definition, planning, enactment, monitoring, and evaluation. Supporting this view, Strømsø & Bråten (2010) found that adaptive Internet-specific justification beliefs, that is, the view that knowledge claims on the Web about a specific topic should be evaluated through independent thinking as well as through the comparison of multiple sources, were related to an increased use of self-regulatory strategies while studying on the Web as reported by the students. Few studies have explored the link between epistemic beliefs and self-regulation as indicated by students' navigation on the Web. In a recent study on this issue, Kammerer and Gerjets (2012) explored the relationship between Internet-specific EBs and navigation while students used Google to learn about a conflicting topic. Results of that study indicated that the belief that the Web contains correct knowledge was related to an increased selection of objective (i.e. scholarly or informational) web pages during web search.

Our study aims at expanding these results by exploring the impact of topic-specific EBs about climate change in the exploration of search engine results pages (SERPs) and source information presented in web pages, navigation across pages, and page bookmarking while studying a controversial topic on the Web. We expected that students with more adaptive topic-specific EBs

would process for a longer time the search results presented in the SERP as well source information provided in the web pages, would visit more web pages, and would bookmark pages for further study based on a more systematic evaluation of page characteristics rather than on link position.

## Method

Thirty undergraduate students received the following instructions: “Imagine you are taking an introductory course on Climate Science. The professor had requested you to work in small groups to write a report on the topic of climate change. Your group has decided to divide up the work into several subtopics, and you will be in charge of the subtopic ‘Reduction of greenhouse gas emissions’. For the next group meeting your task is to search the Internet to look for relevant pages and to bookmark the two pages that cover best your subtopic, so that your team colleagues could have a complete understanding of the topic by only reading those two pages.” We used a mock Google SERP that displayed ten real web pages, including institutional, personal, and commercial sites. The three pages at the top of the SERP were more relevant and complete than pages at lower levels, as rated by independent raters. Participants' eye movements and navigation logs were recorded during the task and their EBs were assessed with the Topic-Specific Epistemic Belief Questionnaire (Bråten et al., 2009).

## Results and discussion

Results are reported for the two EB dimensions concerning the process of knowing (source of knowledge and justification for knowing, cf. Hofer & Pintrich, 1997), because only those were related significantly with several navigation indexes. The 'source of knowledge' dimension assessed the belief that knowledge is constructed by the self (high scores) rather than transmitted by external authorities (low scores). This dimension was not related to either total inspection times of search results in the SERP or total inspection times of source information (i.e. page logos) provided in the web pages visited ( $r(30) = .02, p = .90$ ;  $r(30) = -.03, p = .88$ ; respectively). Regarding page visits, source beliefs positively correlated with the percentage of visits to the first few pages provided in the SERP (pages 1-3, relevant and complete),  $r(30) = .33, p = .06$ , but not with pages displayed further down in the SERP ( $r(30) = .24, p = .19$ ,  $r(30) = .23, p = .22$ , for pages 4-7, i.e. relevant but less complete, and pages 8-10, i.e. irrelevant, respectively). Finally, regarding page bookmarking, we constructed a ‘list index’: each page was given a score indicating their location in the SERP list, 1 to 10 from top to the bottom. The list index was an average of the two pages bookmarked. There was a negative correlation between source beliefs and the bookmarking list index ( $r(30) = -.34, p = .05$ ), indicating that students who belief that knowledge is constructed by the self bookmarked pages more at the top of the SERP. The correlation between percentage of visits to the first pages (pages 1-3) and the bookmarking list index was significant as well ( $r(30) = .46, p = .01$ ), a pattern of results that warranted the potential existence of mediation effects. Thus, we tested for possible mediation effects of source beliefs on bookmarking through navigation (i.e. percentage of visits to initial pages). The results revealed a significant positive indirect effect (estimate: 0.09; CI95%: 0.01 to 0.29). Thus, students who belief that knowledge is constructed by the self visited a higher percentage of pages on top of the SERP and as a consequence of that they bookmarked more often web pages from the top of the SERP.

The 'justification for knowing' dimension assessed the belief that knowledge claims need to be justified through rules of inquiry and cross-checking of different knowledge sources (high scores), rather than on the basis of what feels right (low scores). With regard to these beliefs, results showed significant positive correlations with total inspection times of the search results in the SERP and total inspection times of source information in the web pages ( $r(30) = .43, p = .01$ ;  $r(30) = .39, p = .03$ ; respectively). Results concurred with the claim that students with sophisticated justification for knowing beliefs more often checked the sources to corroborate the web pages' claims. In addition, whereas justification beliefs didn't correlate with percentage of visits to the first pages (pages 1-3,

relevant and complete),  $r(30) = .01$ ,  $p = .94$ , they did so with percentage of visits to pages displayed further down in the SERP ( $r(30) = .38$ ,  $p = .04$ ,  $r(30) = .38$ ,  $p = .04$ , for pages 4-7, i.e. relevant but less complete, and pages 8-10, i.e. irrelevant, respectively). This indicated that students with sophisticated justification beliefs compared more web pages to meet their learning goal, including those displayed further down the SERP. Finally, justification beliefs didn't correlate with the bookmarking list index ( $r(30) = .02$ ,  $p = .90$ ), which reveals that this dimension was not related to the use of a 'Google trust' heuristic when it comes to page bookmarking.

Our study provides new insights on how undergraduates' topic-specific EBs are reflected in how they navigate on the Web and how they explore the web pages visited when they search to address a controversial topic. Specifically, students who belief that knowledge is constructed by the self rather than coming from experts just tend to visit more often the pages on top of the SERP, and as a consequence they bookmark pages from this range. Although following this approach might be reasonable to access simple facts, it might be risky when searching for conflicting topics, because students might end up studying information not necessarily coming from trustworthy pages. On the contrary, students who belief that knowledge claims need to be justified by rules of inquiry and cross-checking of knowledge sources visit a higher range of web pages rather than sticking to the ones at the top of the SERP. They also pay more attention to source information in the webpages visited. As a result they do not just bookmark the top pages provided by the SERP for further study.

From a practical point of view, the findings of the present study may inform the development of concrete interventions how to critically evaluate and use different information sources when searching for controversial topics on the Web. As our findings suggest, such interventions may profitably include an explicit focus on epistemic beliefs.

## References

- Bråten, I., Britt, M.A., Strømsø, H.I., & Rouet, J.F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Towards an integrated model. *Educational Psychologist*, 46, 48-70.
- Bråten, I., Gil, L., Strømsø, H.I., & Vidal-Abarca, E. (2009). Personal epistemology across cultures: Exploring Norwegian and Spanish university students' epistemic beliefs about climate change. *Social Psychology of Education*, 12, 529-560.
- Hofer, B.K., & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
- Kammerer, Y., & Gerjets, P. (2012). Effects of search interface and Internet-specific epistemic beliefs on source evaluations during Web search for medical information: An eye-tracking study. *Behaviour & Information Technology*, 31, 83-97.

# Move Like a Fish: Do Gestures Aid Learning from Photographs and Videos?

*Katharina Scheiter, Jana Arndt, Birgit Imhof*

Knowledge Media Research Center, 72076 Tübingen, Germany  
{k.scheiter, j.arndt, b.imhof}@iwm-kmrc.de

*Shaaron Ainsworth*

School of Psychology, University of Nottingham, United Kingdom  
Shaaron.Ainsworth@nottingham.ac.uk

**Abstract.** Gestures have been shown to foster communication and learning in a variety of tasks. The present study investigated whether gestures would help learning of dynamic behaviour (i.e., fish locomotion behaviour) and whether their effect would be dependent on the format of the visual representations used. The 2x2 between-group design ( $N = 88$  students) resulted from cross-varying visualization format (static versus dynamic) and gesture (with versus without gesture instruction) during the learning phase. To assess learning outcomes, students were asked to classify fish based on their locomotion behaviour as well as to describe fish showing familiar as well as unfamiliar locomotion behaviours. Students' behaviour was video-taped during learning and testing. Initial results showed that instructed as well as spontaneously performed gestures improved performance in the classification task irrespective of visualization format. Data coding for the description tasks and of students' gestures performed during learning is ongoing, and further results will be presented at the conference.

**Keywords:** visualization, gesture, learning about movements

## Introduction

Recently, there has been an explosion of interest in the role of gesture in human cognition. Gestures are not epiphenomena but instead play an integral role in communication; helping not only listeners but also speakers. Consequently, researchers reasoned that if gesturing helps people think, then it may also help them learn and there is evidence for a range of roles for gesture in learning. For example, when teachers gesture to accompany speech, preschool children learnt more (Valenzeno, Alibali, & Klatzky, 2003); and gestures help children remember what they have learnt (both arbitrary word lists and solutions to math problems; Goldin-Meadow & Wagner, 2005). Gestures also benefit learning even when subsequently not allowed: Chu and Kita (2011) found that encouraging participants to gesture during spatial problem solving increased performance on the tasks **and** transferred to spatial visualization problems even when gesturing was prohibited.

So far, little research has explored whether the form of the to-be-learnt materials influences gesturing. De Koning and Tabbers (2011) argue that gestures play four keys roles in learning from dynamic representations; the one explored here is when gestures are used to follow presented movements – in this case movements of fish. However, it is unclear whether gesturing movements will be most useful when learners are presented with dynamic representations or static ones. On the one hand, learners with static pictures may benefit the most. Hegarty (2004) found that gestures are useful when learners need to mentally animate from static pictures as gestures proceed verbal understanding and help learners to offload cognition. In addition, it may be the case that being forced to infer the movements and gesture them will result in more effortful learning (de Koning & Tabbers, 2011), which researchers argue should lead to deeper learning. But, on the other hand, videos have more complete information for learners to copy and so learners may embody movement information more accurately. The present study aims to evaluate these different explanations by assessing learning outcomes, gesture qualities, and mental effort as a function of students being asked to gesture or not whilst watching videos or sequences of photographs of swimming fish.



## Method

Eighty-eight students from the University of Tuebingen, Germany, participated in the study (61 female;  $M = 24.17$  years). The 2x2 between-group design resulted from cross-varying visualization format (static versus dynamic) and gesture (with versus without gesture instruction). They had no prior domain knowledge as determined by a questionnaire administered prior to the study.

The learning materials consisted of visualizations of 6 fish whereby each fish deployed a different locomotion behaviour. Depending on condition, there was either a single, non-interactive video of each fish or a series of 9 sequentially presented photographs (key frames from the videos representing the most crucial steps in each movement, chosen with the help of a domain expert).. Both visualizations automatically restarted after one loop had been shown; learners were asked to press the space bar once they felt confident they had understood the fish's movement behaviour. The sequence of visualizations for the 6 fish was presented twice to students. In the conditions with gesture instruction, prior to learning students were asked to use their hands and/or body to gesture in a way that they thought would help them to better understand the fish's movements. As dependent variables we assessed test performance, students' use of gestures, learning time and mental effort.

In the first test (near transfer description test), students were shown videos of new fish that performed the same movements as the previously studied fish and were asked to describe the movement as precisely as possible. The second test (categorization test) comprised videos of fish displaying the same movements as the previously studied fish. Students classified each new fish by assigning it to the corresponding fish from the learning phase (represented by a static picture). Finally, new videos were shown of fish that showed different movement behaviours. In this far transfer description test students were again asked to describe the fish's movement behaviour as precisely as possible. To assess students' use of gestures participants were video-taped during the whole experiment with two cameras positioned at different viewing angles. The recordings are currently being coded for features including frequency of gestures and involvement of various extremities (e.g., hands, arms, upper body). Mental effort during learning was assessed using a 7-point Likert scale from 'none at all' to 'very much' which students completed after each (first) presentation.

## Results

Four students were excluded for failing to follow instruction or because of prior experience with the topic. A 2x2-ANOVA (visualization format  $\times$  gesture) for classification performance showed a main effect of gesture,  $F(1,84) = 4.81$ ,  $p = .03$ , with students in the conditions with gesture instruction ( $M = 5.18$ ) outperforming those in the conditions without gesture instruction ( $M = 4.47$ ). There was no main effect of visualization format or interaction. Learning time differed between the two visualization conditions,  $F(1,84) = 285.76$ ,  $p < .001$ , as well as between the gesture conditions,  $F(1,84) = 9.76$ ,  $p = .002$ . Finally, an interaction,  $F(1,84) = 4.99$ ,  $p = .03$ , revealed that gestures increased learning time, but that this effect was observable only in the pictures but not in the video conditions. Overall, pictures were studied for a longer time than the videos. However, learning time was unrelated to test performance ( $r = .16$ ). With regard to mental effort, there was a marginally significant main effect of visualization format,  $F(1,84) = 3.49$ ,  $p = .065$ , with students in the pictures conditions reported higher invested mental effort ( $M = 4.96$ ) than those in the video conditions ( $M = 4.38$ ). There was no effect of gesture and no interaction nor was mental effort related to classification performance ( $r = .04$ ).

Contrary to our expectations, we observed that 8 students in the no-gestures condition spontaneously gestured without being so instructed: 3 in the pictures condition (13%) and 5 in the video condition (22.7%). Consequently, an exploratory analysis was conducted within the conditions **without** gesture instruction to compare the 8 students who spontaneously gestured (gesturers) to those 37 students who did not (non-gesturers) irrespective of visualization format. Interestingly, gesturers ( $M = 5.50$ ) outperformed non-gesturers ( $M = 4.24$ ) in the classification test,  $F(1,43) = 4.81$ ,  $p = .03$ ,

while not differing from them in either learning time or mental effort. The positive effect of spontaneous gestures was almost twice as large ( $\eta^2 = .10$ ) as that of instructed gestures ( $\eta^2 = .06$ ).

## Discussion

The present study investigated the effects instruction to gesture whilst learners watched either photos or videos of fish performing different types of locomotion behaviours on students' ability to recognize and accurately describe these locomotion behaviours as well as their perceived mental effort during learning. Contrary to prior studies (Imhof, Scheiter, & Gerjets, 2011), there was no effect of visualization format on learning outcomes. However, learners spent less time studying videos than pictures and reported less mental effort. More importantly, the instruction to gesture improved performance in a classification task (cf., Chu & Kita, 2011). Even though learners only had available the human body parts to gesture complex fish movements, gestures were able to improve students' ability to recognize the different fish locomotion behaviours. Only two students did not gesture at all despite being instructed, suggesting that learners are able to use gestures once instructed. Interestingly, there was also evidence for spontaneous gestures, which even had a stronger effect on learning than instructed ones and compared to the latter did not even increase learning time. This finding suggests that the effects found in the gestures condition are specific to gesturing and not caused by deeper engagement in the task in the general. Nevertheless, future studies need to address this by using control conditions with similar levels of task involvement. Initial observations of the video process data indicate that students vary largely in the extent to which they gesture and how continuously and accurately they do so. It is expected that actual gesture behaviour will explain additional variance regarding test performance. Moreover, stronger gesture effects are expected for the open description tasks, which focus more strongly on students' abilities to retrace the various fish movements. Effects of gesturing of classification accuracy were not moderated by visualization format. However, it might well be that visualization format affected **how** students gesture, with different gesturing approaches being observable for static and dynamic visualization formats.

## References

- Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology-General*, 140, 102-116.
- de Koning, B., & Tabbers, H. (2011). Facilitating understanding of movements in dynamic visualizations: An embodied perspective. *Educational Psychology Review*, 23, 501-521.
- Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *Trends in Cognitive Sciences*, 9, 234-241.
- Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, 8, 280-285.
- Imhof, B., Scheiter, K., & Gerjets, P. (2011). Learning about locomotion patterns from visualizations: Effects of presentation format and realism. *Computers & Education*, 57, 1961-1970.
- Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28, 187-204.

# Information Access Patterns of Students Working on Text-Picture Integration Tasks

*Wolfgang Schnotz<sup>1</sup>, Ulrike Hochpoechler<sup>1</sup>, Thorsten Rasch<sup>1</sup>, Mark Ullrich<sup>2</sup>, Holger Horz<sup>2</sup>, Nele McElvany<sup>3</sup>, Jürgen Baumert<sup>4</sup>*

<sup>1</sup> University of Koblenz-Landau, <sup>2</sup> University of Frankfurt, <sup>3</sup> Technical University of Dortmund, <sup>4</sup> Max-Planck-Institute of Human Development, Germany  
schnotz@uni-landau.de

**Abstract.** When students read for learning, they have to integrate text and graphics information into coherent knowledge structures. This study aimed at analyzing how students deal with texts and pictures when they integrate the two sources of information. Furthermore, it investigates differences between different school types and grades. 40 students from grades 5 and 8 from higher track and lower track schools were asked to integrate texts and graphics in order to answer items from different levels of a text-picture integration taxonomy. Students' eye movements were recorded. Results suggest different functions of text and graphics associated with different processing strategies. Texts are more likely to be used according to a coherence-formation strategy, whereas graphics are more likely to be used on demand as visual cognitive tools by an information selection strategy. Students from different tracks of schooling revealed different adaptivity with regard to the requirements of combining text and graphic information.

## Objectives

Students in secondary school are usually expected to learn from text combined with instructional graphics such as schematic diagrams, maps and various kinds of graphs that visualize and explain complex subject matter. When learning from such materials, students are expected to use text and graphics as complementary sources of information, because they have to integrate verbal and pictorial information for the construction of mental representations of the learning content (Mayer, 2001). The following research questions arise: Which strategies are used when students process text information and which strategies are used when they process picture information? Which differences exist regarding strategy usage between students from different grades and from different types of schools?

## Theoretical Framework

The model of integrated text and picture comprehension (Schnotz & Bannert, 2003) was used as a theoretical basis for answering these questions. The model assumes that text and picture integration takes place in a working memory, which has a limited capacity and includes different channels for processing verbal and pictorial information. Integrating verbal and pictorial information requires mappings between corresponding elements in the text and the picture. Surface structure mappings include interrelations between words and graphical elements based on cohesive devices such as color coding, common numbers, symbols or labels. Deep structure mappings include interrelations between conceptual structures and mental models of the learning content. Within deep structure mapping, a distinction can be made between different taxonomy levels of extracting information from text and pictures: Level A refers to extracting detail-information, level B refers to extracting simple relations, and levels C refers to extracting complex relations.

## Method and Data Sources

We presented four text-graphic combinations to students from different grades and from different tracks of schooling and asked them to answer sequentially presented items from different taxonomy levels. Each text-graphic combination was combined with a sequence of three items: The first item was from taxonomy level A, the second item was from taxonomy level B, and the third item was from taxonomy level C. The text-graphic combinations were selected from a larger sample of 48 text-picture integration units that were taken from textbooks of biology and geography between grade 5

and grade 8. 20 students from the higher track of the German school system (Gymnasium) and 20 students from the lower track (Hauptschule) participated in the study. In each group, half of the students were from grade 5 and the other half were from grade 8.

The students' eye movements during reading the text and observing the pictures were registered by a head mounted EyeLink II tracker. For each unit and each item, the total fixation times were determined both for the text, for the picture, and for the item. For the following analysis, total fixation times on the text were aggregated for answering the level A items across all four units, for the level B items across all four units, and for the level C items across all four units. Similarly, total fixation times for the picture(s) were aggregated across all four units for the level A, the level B, and the level C items.

## Results

The ANOVA with the factors 'item' (level A/level B/level C), 'school type' (higher track/lower track) and grade (5th/8th) for the total fixation times of the text revealed a significant main effect of the factor 'item' and a significant interaction 'item x school type'. The sharp decrease of text fixation times from item A to item B indicates that there is an intensive and time-consuming phase of initial mental model construction, before the first item can be answered. After answering the first item, the text usage for answering the second item decreases dramatically. After the initial mental model construction, the text seems to be used only for some item-specific mental model updates, although the item difficulty continuously increases. The decrease is different between the school types: Students from the lower track seem to have more difficulties during the initial mental model construction than students from the higher track.

The ANOVA with the factors 'item', 'school type' (higher track/lower track) and grade (5th/8th) for the total fixation times of the graphic revealed a significant main effect of the factor 'item', a significant main effect of the factor 'school type' and a significant interaction 'item x school type'. The result pattern for the fixation times of the graphics across the three levels of items differs remarkably from the fixation times of the text. Besides the fact that time on text was considerably higher than the time on graphics, the fixation times of the text decreased continuously, whereas the fixation times of the graphics first decreases from the A item to the B item, but then increases from the B to the C item.

## Summary and Discussion

The findings of this study indicate that text and graphics play fundamentally different roles in mental model construction, even when they constitute a well-formed coherent set of verbal and pictorial information. The dynamic of mental model construction consists of an extensive use of text for initial mental model construction before any question answering, whereas graphics are used more ad-hoc as external visual tool for the questions at hand. This implies that processing of text is less item-dependent and insofar more a bottom-up process, whereas processing of graphics is more item dependent and insofar more a top-down process. Top-down processing requires strategies, and it seems that the lower track students frequently lack these strategies, because they are less able to adapt their processing to the requirements of the task at hand.

As for students from the higher track, text and graphics seem to serve different functions and are therefore used according to different strategies. Texts are more likely to be used according to a coherence-formation strategy than graphics: Texts guide the reader's conceptual analysis by a description of the subject matter, which results in a coherent semantic network and mental model regardless of the difficulty of the item at hand. Graphics are more likely to be used according to a task-specific information selection strategy than texts. Graphics serve as scaffolds for initial mental model

construction, but are afterwards more likely to be used on demand as easily accessible external representations for item-specific mental model updates.

As for students from the lower track, the situation is somewhat different. Texts are also more likely to be used by these learners according to a coherence-formation strategy. They obviously have more difficulties with word recognition and lexical access, which is indicated by their average fixation times. However, they nevertheless invest a high amount of time into the text during the first phase of initial mental model construction, although the item to be answered is relatively easy. Afterwards, they invest much less time into the text, even when the following items are more demanding. Graphics seem also to serve for lower track students as scaffolds for initial mental model construction. Contrary to higher track students, however, the lower track students do not use graphics afterwards more intensively when items become more difficult. Contrary to our expectations, lower track students (as well as younger students) seem not to use graphics more intensively and text less intensively than higher track students (and older students). However, the results are in line with results of Hannus and Hyönä (1999), who found that 10 year-old children's learning was heavily driven by the text and that children inspected illustrations only superficially.

Whereas the development of text reading skills receives justifiably much attention in schooling, reading graphs and other types of visuals is only briefly addressed in most school curricula. Integrative processing of text and pictures puts specific demands on the learner, because specific strategies of how to employ the mutual constraints between texts and pictures are required. In the present forms of schooling, the development of these skills seems to be a by-product rather than a result of systematic teaching and learning, although integrative processing of texts and pictures is becoming an increasingly fundamental requirement in everyday life and further education.

## References

- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, 24, 95-123.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representations. *Learning and Instruction*, 13, 141-156.

# Unwinding the Relationship between Cognitive Processes and Gaze Behaviour during Multimedia Learning

*Carina Schubert*

*Katharina Scheiter*

*Anne Schüller*

Knowledge Media Research Center  
Schleichstraße 6, 72076 Tübingen, Germany  
{c.schubert; k.scheiter; a.schueler}@iwm-kmrc.de

**Abstract.** Although adding visualizations to text generally improves learning, this *multimedia effect* is not always found. This might be caused by learners' failure to process the materials adequately. This project aims at designing an adaptive multimedia system that fosters strategic use of the learning material. We want to use eye tracking as an online measure for learning difficulties, providing additional information (zoom-ins, highlights) tailored to these difficulties. The objective of the present study is to identify those eye tracking parameters (e.g., long dwell times or rereading of sentences) that indicate comprehension problems. For this purpose, participants learn with a static multimedia learning system and their gaze movements are recorded. Afterwards, verbal data on participants' thoughts during learning are collected using *cued retrospective reporting*. Combination of the data from verbal protocols and learning outcome measures will be used to assess comprehension problems. Correlations between eye-movement parameters and comprehension problems will be explored.

**Keywords:** multimedia learning, eye tracking, cognitive processes, verbal protocols

## Introduction

Research in the last 30 years has yielded a vast amount of evidence for the benefits of combining text with visualizations (for an overview, see Anglin, Vaez, & Cunningham, 2004; Mayer, 2009). However, learning outcome is not always enhanced by multimedia material (cf. Levie & Lentz, 1982, for an overview). Beyond an inappropriate design of the materials, learners may inadequately process the instructional material, thereby failing to learn from it. Accordingly, one way to support multimedia learning may be to design the learning system in a way that fosters strategic use of the material. Thus, the aim of this project is to design an adaptive multimedia learning environment that uses learners' gaze behaviour as an online measure for learning difficulties and provides support tailored to these difficulties. Specifically, the learner will be presented with additional, supportive information (e.g., highlights, zoom-ins) when eye tracking parameters indicate that he/she struggles to understand something. The design of such an adaptive system requires a systematic assessment of how to interpret eye tracking parameters during multimedia learning. Accordingly, the objective of the present study is to gain insight into the relationship between eye tracking parameters, cognitive processes and learning outcomes.

Eye tracking is a method that has been widely used to gain insight into cognitive processes during reading (for an overview, see Rayner, 1998). It can provide useful information on the spatial and temporal allocation of visual attention (Rayner, 1998). Nevertheless, there has so far been little investigation of the processing of multimedia learning material using eye tracking (Hyönä, 2010). One major problem with the interpretation of eye movements is that the validity of their interpretation as a measure for cognitive processes is limited. Gaze parameters like dwell time or total fixation duration can tell us where and when a learner allocated visual attention, which external representation the learner regarded or how often he/she switched between representations. However, these parameters do not tell us why a learner spends time on a text passage or on a picture. For example, a long dwell time might indicate deeper processing of the learning content for a learner with high prior knowledge but comprehension problems for a learner with low prior knowledge (Schwonke, Berthold, & Renkl, 2009). Furthermore, gaze data alone is no indicator for the success of understanding a piece of

information (Hyönä, 2010). To make the information extracted from gaze data less ambiguous, they must be supplemented with additional data like think-aloud protocols or learning outcome in a first step. A technique that has been shown to produce rich data and may act as a reliable source of insight into students' cognitive processing is *cued retrospective reporting* (Van Gog, Paas, Merriënboer, & Witte, 2005). Thereby, following the learning phase, participants are presented with a video showing their gaze movements superimposed onto the learning material (i.e., a gaze replay) and asked to describe what they thought during learning from the material. By combining eye tracking and cued retrospective reporting, we want to assess how to interpret gaze parameters during multimedia learning. In the long run, we want to be able to identify eye tracking parameters as online measure of comprehension problems. More specifically, they will be the starting point for the design of an adaptive multimedia learning system that automatically adapts to each learner's gaze behaviour.

## **Method**

### *Goal*

The present study aims at investigating the multimedia learning process by collecting gaze data and cued retrospective reporting. We also want to take learners' prior knowledge into account, assuming that gaze parameters need to be interpreted differently for learners with varied levels of prior knowledge. We are especially interested in the diagnosis of comprehension problems during learning.

### *Materials and Procedure*

For this purpose, we first assess several learner variables such as age, gender, field of study, last mark in biology and self-reported interest on biology as control variables. After that, domain-specific prior knowledge is tested with 25 multiple-choice questions on scientific literacy and 15 multiple-choice questions on cell division.

Learners are then presented with a computer-based multimedia learning system on mitosis. The learning material consists of ten pages with text on the left and a corresponding static visualization on the right. Learners can proceed at their own pace. During learning, their gaze movements are recorded with an SMI RED250 remote eye tracker.

They then complete a posttest consisting of 16 multiple-choice questions (text- and picture-based), a free recall question and a fill-in-the-blank text. The posttest assesses recall as well as transfer.

Afterwards, we collect verbal data via cued retrospective reporting, providing qualitative data on each participant's cognitive processes during the learning phase.

### *Analysis*

Combination of the information from verbal protocols and analysis of the learning outcome measures will allow for getting a detailed picture of where a learner had comprehension problems. Therefore, we will develop a coding scheme for the verbal protocols to identify at which points during learning each participant struggled to comprehend the learning content. It will then be explored which eye-movement parameters correlate with comprehension problems for learners with different levels of prior knowledge.

## **Preliminary Results and Discussion**

A first analysis of verbal and gaze data from 32 participants suggests that a good indicator of comprehension problems during learning is the rereading of sentences or paragraphs: participants frequently verbalized that they did not know a concept or did not understand a sentence and therefore reread it. In contrast, dwell times presumably cannot be interpreted independently of a learner's prior

knowledge level. For learners with low prior knowledge, they can indicate comprehension problems, whereas for learners with high prior knowledge, they point to deeper processing of the content.

Data collection is currently running; by the time of the conference, we will have collected and analyzed data from 60 participants.

## References

- Anglin, G. J., Vaez, H., & Cunningham, K. L. (2004). Visual representations and learning: The role of static and animated graphics. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 865–916). Mahwah, NJ: Erlbaum.
- Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*, 172-176.
- Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology Journal, 30*, 195-232.
- Mayer, R. E. (2009). *Multimedia learning (2nd ed.)*. New York: Cambridge University Press.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.
- Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology, 23*, 1227-1243.
- Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*, 237-244.



# What Pictures Are Good For

*Anne Schüller<sup>1</sup>, Katharina Scheiter<sup>1</sup>, Francesca Pazzaglia<sup>2</sup>*

<sup>1</sup>Knowledge Media Research Center  
Schleichstraße 6, 72076 Tübingen, Germany  
{a.schueler, k.scheiter}@iwm-kmrc.de

<sup>2</sup>University of Padova  
Via Venezia 8, 35131 Padova, Italy  
francesca.pazzaglia@unipd.it

**Abstract.** The aim of the reported study was to test whether pictures are especially helpful to convey spatial information, but less helpful to convey non-spatial (e.g., visual) information. Ninety-nine students learned in five different learning conditions (only picture, picture plus text with either visual or spatial information, only text with either visual or spatial information). Afterwards they had to verify verbal statements asking for either spatial or visual information (within-subject factor). As expected, pictures supported learning of spatial information but not learning of visual information. Interestingly, presenting pictures together with text containing spatial information even reduced learning outcomes. The implications of these findings for explaining the multimedia effect are discussed.

**Keywords:** multimedia effect, spatial information, visual information

## Introduction

In the last decades there has been a lot of research regarding the effects of adding pictures to text. A large body of evidence supports the assumption that learners benefit from multimedia presentations consisting of text and pictures (for an overview see Mayer, 2009). According to the process-oriented view (Schmidt-Weigand & Scheiter, 2011) this multimedia effect can among other things be explained by the fact that pictures facilitate spatial reasoning (Larkin & Simon, 1987). Therefore, pictures should support learning especially when spatial cognitive processes are required. However, based on the process-oriented view one might assume that presenting pictures alone might already be sufficient to solve spatial reasoning tasks (e.g., Meneghetti, Borella, Grasso, & De Beni, 2012), because the required information can be extracted from the picture alone. Indeed, there is evidence that verbal information can even hinder learning from pictures when spatial information has to be conveyed (Schmidt-Weigand & Scheiter, 2011). Therefore, one aim of the reported study was to test whether presenting pictures alone might be equally or even more helpful than presenting text-picture combinations or texts alone when learners have to solve spatial tasks. On the other hand, the process-oriented view suggests that pictures should be less helpful for learning when non-spatial information has to be learned (e.g. visual information like color or size, see Kulhavy et al., 1993). Therefore, the second aim of the reported study was to test whether presenting text alone might be as helpful as presenting text with pictures and even more helpful than presenting pictures alone when learners have to solve tasks requiring visual memory and inferences.

## Method

Ninety-nine students (72 female, average age:  $M = 23.47$ ,  $SD = 3.22$ ) participated in the experiment, which was based on a  $5 \times 2$  mixed design with learning condition (description see below) as between subject factor and information type (visual vs. spatial information) as within-subject variable.

All materials were presented on a computer screen. They comprised an introduction, the system-paced learning phase, and a test phase. Learning condition was varied in the learning phase as follows: learners received only pictures, only texts containing visual information, only texts containing spatial information, pictures together with texts containing visual information, or pictures together with texts containing spatial information. The two texts containing spatial contents were adapted from Pazzaglia et al. (1994) and described spatial relations within a tourist center and a holiday farm (e.g., North of

the lake is the highway, which stretches from east to west above the entire area). The texts containing visual information (e.g., The highway has a grey color and is lined by firs which are of olive-green color) resembled the spatial text versions with respect to the Flesch reading ease score (Flesch, 1948) and the number of propositions therein. The pictures depicted the spatial relationships as well as visual information of the two locations (i.e., size, shape, and color of the landmarks). In the picture-only condition, the landmarks in the pictures were labeled (e.g., restaurant, farmer house) since the verbal test items referred to these objects, therefore requiring unambiguous reference.

After the learning phase, learners had to verify first verbal recall and then inference statements (the latter similar to those used by Pazzaglia et al., 1994) that referred to either visual (e.g., “The hotel has the same color as the rehabilitation center”) or spatial (e.g., “The hotel is in the north-east of the rehabilitation center”) information presented in the learning phase. The visual and spatial statements were presented in random order. Note that depending on experimental condition the information necessary to answer the verbal statements was presented in the text, the picture or in both.

## Results

To test our assumptions, a mixed MANOVA with learning condition as between-subject factor, information type as within-subject factor and recall and inference performance as dependent variables was conducted. The MANOVA revealed significant main effects of learning condition and information type as well as a significant interaction between both factors, all  $ps < .001$ .

The follow-up ANOVAS showed for recall as well as inference performance main effects of learning condition and information type, which were – for both dependent variables – qualified by the disordinal interactions depicted in Figure 1<sup>9</sup>.

Bonferroni-adjusted comparisons showed that for spatial recall and spatial inference statements learners who received only pictures outperformed learners who received only text containing spatial information and even learners who received pictures together with spatial or visual text information. Furthermore, no multimedia effect appeared, that is, learners with pictures accompanied by text containing spatial information did not outperform learners who received only text containing spatial information. This pattern of results speaks clearly in favor of the hypothesis that pictures are especially suited to convey spatial information and that the presentation of additional verbal information does not facilitate the extraction of these kinds of information, but may even hinder it.

With regard to visual recall and visual inference statements the pattern of results changed completely: Regarding recall performance no significant differences were observed. For inference performance, however, Bonferroni-adjusted comparisons showed that learners who received pictures plus text containing visual information outperformed learners with pictures plus text containing spatial information as well as learners who received only pictures. On the other hand, learners with pictures and texts containing visual information and learners who received only text containing visual information showed the same level of performance. This pattern of results speaks clearly in favor of the hypothesis that the visual information is not extracted from the picture but from text. However, presenting a picture additionally did not hinder the extraction process.

---

<sup>9</sup> Note, that learners who received only text containing visual information performed at chance level regarding spatial statements and learners who received only text containing spatial information performed at chance level regarding visual statements. Therefore, they performed worse than learners in the other conditions which we will not discuss further.

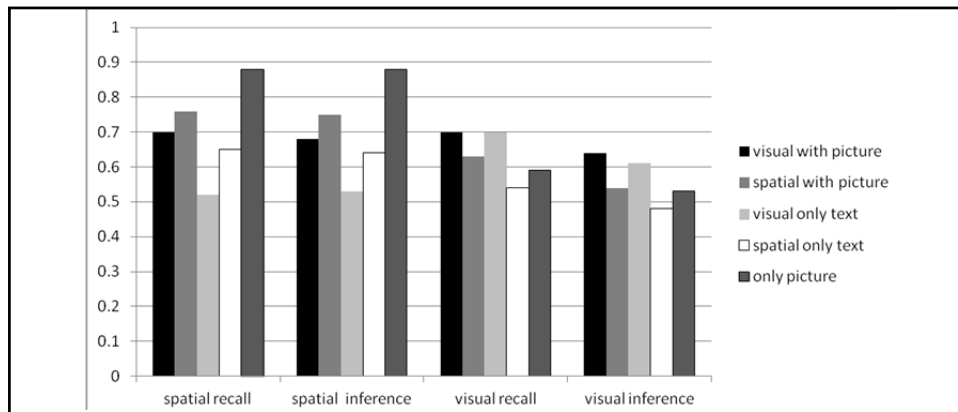


Figure 1. Spatial recall and inference (left) and visual recall and inference (right) as a function of learning condition.

## Discussion

In line with the process-oriented view, the results demonstrate that pictures aid learning when spatial information has to be conveyed. Interestingly, this beneficial effect of picture presentation disappeared when texts were presented in addition. On the other hand, pictures did not aid learning when visual information had to be conveyed. Here, texts proved to be much better suited. Overall, these results indicate that text and pictures serve different learning goals: whereas pictures support the acquisition of spatial information, texts support the acquisition of visual information. Thus, the multimedia effect might be explained by the fact that normally post tests consist of both, items asking for spatial and visual information. To solve such post tests, learners can rely on pictures to answer the spatial items whereas they can rely on texts to answer the visual items. In sum, by combining text and pictures learners may profit from representations, which in turn might explain the multimedia effect. Future studies should investigate whether text-picture overlap (i.e., redundancy) or text-picture contiguity (e.g., integrated format) may influence the observed results. Moreover, individual differences like spatial abilities or gender should be taken into account.

## References

- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Kulhavy, R. W., Stock, W. A., Verdi, M. P., Rittschof, K. A., & Savenye, W. (1993). Why maps improve memory for text: the influence of structural information on working memory operations. *European Journal of Cognitive Psychology*, 5, 375-392.
- Mayer, R. E. (2009). *Multimedia learning. Second edition*. Cambridge: University Press.
- Meneghetti, C., Borella, E., Grasso, I., & De Beni, R., (2012). Learning a route using a map and/or description in young and older adults. *Journal of Cognitive Psychology*, 24, 165-178.
- Paivio, A. (1991). Dual coding theory: Retrospect and current state. *Canadian Journal of Psychology*, 45, 255-287.
- Pazzaglia, F., Cornoldi, C., & Longoni, A. M (1994). Limiti di memoria e specificita' di rappresentazione nel ricordo di descrizioni spaziali "dall'alto" ed "entro il percorso" [Limits of memory and the specificity of representations in the memory of survey and route descriptions]. *Giornale Italiano di Psicologia*, 21, 267-286.
- Schmidt-Weigand, F., & Scheiter, K. (2011). The role of spatial descriptions in learning from multimedia. *Computers in Human Behavior*, 27, 22-38.

# Long-term Multimedia Learning

*Judith Schweppe*

University of Erfurt

P. O. Box 900 221, 99105 Erfurt, Germany

judith.schweppe@uni-erfurt.de

*Ralf Rummer*

University of Erfurt

P. O. Box 900 221, 99105 Erfurt, Germany

ralf.rummer@uni-erfurt.de

**Abstract.** Cognitive theories of multimedia learning provide instructors with guidelines for presenting materials in a manner best fit to enhance learning outcomes. These design principles share the recommendation to facilitate processing. One such principle is the modality principle, according to which pictures should be presented with auditory rather than visual texts. Research on desirable difficulties, however, indicates that – in contrast to short-term learning – long-term learning benefits when learning conditions are more difficult. From this perspective, the question arises whether following multimedia design principles serves short-term but not long-term learning. In Experiment 1, we tested this in terms of the modality principle by varying text modality and time of testing (immediate vs. after one week). Experiment 2 replicated the delayed condition. The findings support the assumption that the combined processing of written text and animations, which is regarded as more difficult than processing of oral text plus animations, improves long-term learning.

**Keywords:** multimedia learning, modality effect, desirable difficulties, long-term learning

The primary goal of research on multimedia learning is to provide instructors with ways of presenting learning materials that are best fit to enhance students' learning outcomes. This has inspired several principles that provide guidelines for the design of multimedia learning materials. The common core of most of these principles is to facilitate processing by reducing extraneous cognitive load in working memory. However, research on episodic long-term memory suggests a different approach: It has been demonstrated repeatedly that good initial performance does not translate reliably into good long-term learning. Instead, learning under cognitively demanding conditions that lead to worse initial performance results in enhanced performance after a delay of one day or longer (*desirable difficulties*, Bjork, 1994). We applied this long-term memory perspective to multimedia learning, which has almost exclusively been investigated with immediate tests. If the conditions that improve short-term learning lead to poorer performance in delayed tests – as is the case with desirable difficulties – the design principles might need to be reversed (for long-term learning).

There are only two studies that tested a design principle, here the modality principle, after a delay (Segers, Verhoeven, & Hulstijn-Hendrikse, 2008; Witteman & Segers, 2010). In their study with children, Segers et al. (2008) indeed observed a reversed modality effect in a test one week after the instruction. However, this inverse modality effect does not suffice to question the applicability of the modality principle to long-term learning, since presentation of the materials was self-paced, which often results in visual advantages even with immediate tests (e.g., Tabbers, 2002).

We tested the modality principle with experimenter-paced learning when tests were administered immediately after the instructional lesson or one week later. According to the modality principle, there should be an advantage of auditory over visual texts for both immediate and delayed tests. According to the desirable difficulties approach, however, visual text presentation should lead to superior performance in the delayed condition.

## Experiment 1

### Method

48 undergraduates from different majors at the University of Erfurt participated in the experiment, which followed a 2 (text modality: auditory vs. visual) x 2 (test delay: 5 minutes vs. 1 week) between-subjects design. Performance on a retention test and a transfer test served as dependent variables.

We used Moreno and Mayer's (1999, simultaneous presentation) original study materials comprised of a prior knowledge checklist, a multimedia presentation of a lesson on lightning formation, and retention and transfer questions. The experiment took place in a classroom setting. Participants in the immediate testing condition completed the retention and transfer tests after viewing the presentation. Participants in the delayed testing condition were tested one week later.

### Results & Discussion

A 2 x 2 between-subjects ANOVA on the retention data revealed no significant effects. Descriptively, there was an auditory advantage for immediate testing (Means: 5.25 vs. 3.25) and no modality difference in the delayed condition (Means: 5.08 vs. 5.25). In the case of the transfer test, there was no main effect for modality ( $F(1,44) = .01, p = .916$ ) but a significant main effect for delay ( $F(1,44) = 6.97; p = .011; \eta^2 = .137$ ). This was, however, qualified by the interaction between modality and delay ( $F(1,44) = 10.71; p = .002; \eta^2 = .196$ ). While immediate transfer performance was better following auditory text presentation as opposed to visual text (planned comparison:  $F(1,44) = 5.02; p = .03; \eta^2 = .171$ ), delayed transfer performance was superior for participants in the visual conditions (planned comparison:  $F(1,44) = 5.71; p = .021; \eta^2 = .225$ ). Figure 1 illustrates the findings.

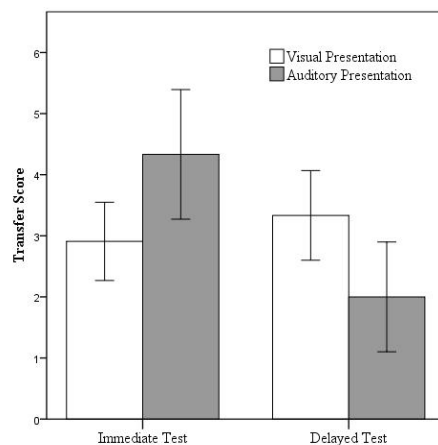


Figure 1. Means for transfer test scores in Experiment 1 as a function of text modality (visual vs. auditory) and time of testing (immediate vs. delayed). Error bars indicate 95% confidence intervals.

Our findings reveal different learning outcomes depending on the delay of testing: Only when performance was tested immediately (as is usually the case in experiments on multimedia learning), participants in the auditory group significantly outperformed those in the visual group on the transfer test. When participants were tested one week later, the modality difference for the transfer test reversed. However, potential caveats are the classroom setting and the rather small sample. Therefore, we replicated the critical delayed condition in a laboratory setting with more participants.

## Experiment 2

### Method

64 undergraduates from different majors at the University of Erfurt participated in the experiment. The experiment followed a one-factorial design, in which text modality (auditory vs. visual) was varied between subjects. Performance on a retention test and a transfer test one week after presentation served as dependent variables. Participants were tested in individual sessions. Otherwise, materials and procedure were analogous to the delayed conditions of Experiment 1.

### Results & Discussion

A *t*-test for independent samples on the retention data revealed a significant main effect for modality,  $t(56) = 2.19$ ;  $p = .032$ ;  $\eta_p^2 = .079$ . In this experiment with more statistical power and more experimental control, participants in the visual condition outperformed those in the auditory condition (Means: 5.93 vs. 4.39). For transfer scores there was also a significant main effect of modality,  $t(56) = 2.54$ ;  $p = .014$ ;  $\eta_p^2 = .103$ , with better transfer performance for written than for oral text (Means: 4.59 vs. 3.39).

### General Discussion

Our results cast doubt on the assumption that listening to texts improves *long-term* learning. Instead, they provide evidence for the assumption that combining graphics with written text serves as a desirable difficulty that impedes initial processing and short-term learning but leads to better learning in the long run. Clearly, our findings do not imply that one should design only difficult learning materials. Given that processing capacities are limited, one cannot process everything in depth. When the materials are too difficult, processing will break down. It is thus crucial to determine and predict when difficulties are desirable and when they are undesirable. It is also still important to design materials that avoid cognitive *overload*. Without doubt, the design principles are relevant here. If the learner was unable to comprehend the materials in the first place (due to overly complex contents or an overly demanding design), there is nothing (sensible) that can be consolidated. However, once overload is avoided, it does not seem to be advisable to avoid extraneous cognitive load. Here, the theories of multimedia learning should be extended. Most research in this domain has concentrated on encoding. Since learning is mainly seen as a function of comprehension, it is only natural that working memory as well as how it can be used to its full capacity is the focus. Based on the current findings, the focus should be shifted to long-term memory and processes of consolidation and retrieval.

### References

- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology*, 22, 375-387.
- Tabbers, H. K. (2002). *The modality of text in multimedia instructions: Refining the design guidelines*. Open University of the Netherlands, Heerlen: Unpublished doctoral dissertation.
- Witteman, M. & Segers, E. (2010). The modality effect tested in children in a user-paced multimedia environment. *Journal of Computer Assisted Learning*, 26, 132-142.

# Visual Attention Guidance in Narrated Animations: Understanding Is More than Just Seeing

*Irene T. Skuballa, Rolf Schwonke, Alexander Renkl*

University of Freiburg, Department of Educational and Developmental Psychology  
Engelbergerstr. 41, 79085 Freiburg, Germany  
skuballa@psychologie.uni-freiburg.de

**Abstract.** Narrated animations present information in a transitory and simultaneous way, which causes some learning disadvantages. Coping with these disadvantages might be especially difficult for learners with restricted working memory capacities. Therefore, the effects of support procedures such as cueing or verbal instruction might be moderated by working memory capacity. In our study, the participants ( $N = 66$ ) were randomly assigned to a cueing group, a verbal instruction group, and a no support group. Results showed that the unsupported group displayed the most difficulties in referencing between auditory and visual information, but showed overall the best learning outcomes. Furthermore, instruction impeded learning in low working memory learners, whereas cueing impeded learning in high working memory learners. The cueing group devoted most visual attention to “narration-congruent” areas in the animation, which fostered learning. However, as cueing was not effective in general, there must also be disadvantages of cueing, which are further discussed.

**Keywords:** eye movement, visual attention, animation, working memory capacity

Animations can be useful in demonstrating how systems are composed and how their components work together. However, animations also have disadvantages. Learning can be hindered as information is presented in a transitory way and often simultaneous changes in the visual display have to be processed (Lowe, 1999). To counteract these disadvantages, support procedures have been developed and researched. Cueing and verbal instructions are two possible ways of directing the learners’ visual attention to the right piece of information at the right time. Both methods aim at emphasizing the key information by reducing the search space as well as releasing learners from extraneous information (Lowe & Schnotz, 2008).

Cueing, also called signaling, refers to additional design elements that are assigned to direct visual attention by accentuating relevant parts of an animation. Graphical devices, such as arrows, spreading colors, or boxes are used to cue the information that is most relevant in the immediate moment during the course of the animation. A recently introduced alternative is ‘anti-cueing’ by changing the luminance or acuity of non-relevant parts (Jarodzka et al., 2012). Some findings indicate that cueing during animations can lead to better learning outcomes; however, replications do not always confirm these results (De Koning, Tabbers, Rikers, & Paas, 2010).

While cueing is based on changes in the design, verbal instructions accentuate a rather strategic approach by providing information prior to learning. In order to prevent cognitive overload, learners are verbally instructed to control their attention by focusing on subcomponents of a task. Instruction, specifically the emphasis change approach, manipulates the temporary priority of selected subcomponents while the “whole” stays unchanged and a holistic view is preserved (Gopher, Weil, & Siegel, 1989). This method was successfully employed in the context of video games: Players with instruction outperformed players without such instruction in a Space Fortress game. We pose the question whether these results can be transferred to learning from animations.

Individual differences in working memory capacity (WMC) might determine the necessity or use of such support procedures. In addition, the procedure of verbal instruction might itself contribute to working memory load, because the instruction has to be kept in mind while watching the animation. Cues, on the other hand, in form of blurring, probably should not enhance load but might interfere with spontaneous, holistic processing.

Against this background, we addressed the following hypotheses: (1) Both supported groups, that is, a group with cueing and a group with verbal instruction, perceive less difficulties in relating the animated display and the narration than the unsupported group. (2) (a) Both supported groups devote more visual attention to the narration-congruent parts of the animation. (b) The amount of this attention is positively related to learning outcomes. (3) Both supported groups achieve better learning outcomes. (4) Do individual differences in WMC moderate the effectiveness of the support procedures (aptitude-treatment interaction) (two-sided hypothesis)?

## Method

Participants were 66 students from the University of Freiburg (cueing:  $n = 22$ ; verbal instruction:  $n = 22$ ; no support:  $n = 22$ ). Eye movements were recorded with a temporal resolution of 120 Hz (SMI RED eye tracking system). The part of the animation that was temporarily in the focus of the narration (i.e., narration-congruent) was defined as area of interest (AOI).

The narrated animation covered the functional principles of a parabolic trough power plant. In the cueing condition all parts that were not in the focus of the narration were blurred. In the verbal instruction condition participants read an instruction before the presentation of the animation. They were informed that several things would happen simultaneously and that it was thus crucial to follow the narration and map it to the visual display. Furthermore, they were asked to ignore narration-incongruent events. In the no support condition the animation was presented without any alterations or additional instructions.

WMC was measured by the Letter-Number Sequencing test (WAIS-III). Prior knowledge was assessed by 30 items (Cronbach's  $\alpha = .84$ ). Learning outcomes were measured by a posttest comprising 40 items on the functions, processes, and structures of the system (Cronbach's  $\alpha = .90$ ). Furthermore, we assessed the subjective evaluation of difficulties in making references between narration and the visual display.

Participants were randomly assigned to one of three conditions: cueing, verbal instruction or no support. Before watching the animation, participants completed the pretest. Finally, participants reported their subjective evaluations and worked on the posttest assessing learning outcomes.

## Results

The conditions differed significantly with respect to the perceived difficulties in referencing between visual display and narration,  $H(2)=6.46$ ,  $p = .039$ . The unsupported group ( $M = 5.18$ ,  $SD = 2.46$ ) reported significantly more difficulties than the cueing group ( $M = 3.45$ ,  $SD = 1.57$ ;  $U = 143$ ,  $p = .009$  (one-tailed),  $r = -.36$ ) and more than the verbal instruction group ( $M = 3.82$ ,  $SD = 1.82$ ;  $U = 165$ ,  $p = .033$  (one-tailed),  $r = -.28$ ). Hence, hypothesis 1 was confirmed.

With respect to visual attention allocation on the narration-congruent parts of the animation, we found a significant condition effect,  $F(2, 63) = 11.275$ ,  $MSE = 7,514,224,135$ ,  $p < .001$ , partial  $\eta^2 = .26$ . The cueing group ( $M = 170,518$  msec,  $SD = 26,172$  msec) spent significantly more time on narration-congruent parts than the unsupported group ( $M = 136,243$  msec,  $SD = 25,527$  msec),  $t(63) = 4.404$ ,  $p < .001$ , partial  $\eta^2 = .24$ . However, there was no significant difference between the verbal instruction group ( $M = 141,399$  msec,  $SD = 25,742$  msec) and the unsupported group,  $t(63) = .662$ ,  $p = .510$ , partial  $\eta^2 = .01$ . Hence, hypothesis 2a was only confirmed for the cueing group.

Dwell time on narration-congruent areas was significantly related to learning outcomes,  $r = .389$ ,  $p = .001$ , and  $r = .559$ ,  $p < .001$ , when controlling for condition. Hypothesis 2b was confirmed.

For testing the hypotheses 3 and 4, we performed a regression from learning outcomes on condition and WMC, including the product term of the two predictors in order to test aptitude-treatment interaction effects. There was a significant condition effect on learning outcomes,  $F(2, 60) = 7.047$ ,  $MSE = 298$ ,  $p = .002$ , partial  $\eta^2 = .19$ . Contrary to our expectations, the unsupported group ( $M = 26.59$ ,



$SE = 1.40$ ) outperformed both supported groups (instruction:  $M = 21.31$ ,  $SE = 1.39$ ; cueing:  $M = 22.80$ ,  $SE = 1.44$ ). However, this main effect had to be qualified by a significant interaction between condition and WMC,  $F(2, 60) = 6.001$ ,  $MSE = 253$ ,  $p = .004$ , partial  $\eta^2 = .17$ . Learners with low WMC were impeded by verbal instructions and benefited from cueing. Learners with high WMC were impeded by cueing and benefited from verbal instructions. Taken together, Hypothesis 3 (i.e. effects of support) had to be rejected, whereas concerning hypothesis 4, an interaction between condition and WMC was discovered

## Discussion

The present results seem to be somewhat contradictory at first glance: Cueing enhances attention on narration-congruent parts of the animation, which, in turn, positively influences learning outcomes. Nevertheless, the cueing condition did not outperform the unsupported condition.

An important open question is why cueing does not lead to generally superior learning, although it enhances the dwell time on AOIs which is positively related with learning outcomes. One possibility is that cueing is visually intrusive and unfamiliar, and interfered with the learners' spontaneous holistic processing strategies. This interpretation is in line with the finding that learners with higher WMC—who are supposed to process more information simultaneously—are impeded by cueing. These learners might have enough capacity to process not only the narration-congruent information, but also more holistic aspects of the system. Moreover, cueing might have conveyed a false sense of 'easiness' and thus prevented learners from investing more effort in processing. In the present study, WMC was measured by a verbal task corresponding to the narrative component of the learning environment. Future research should investigate whether these results apply to a visual WMC task—corresponding to the visual display—as well.

In sum, the present study makes the following important contributions to the present literature on learning from animations: (1) It provides further evidence that cueing in animations influences visual attention, but not necessarily fosters learning; the mechanisms underlying these effects should be investigated in further research. (2) WMC should be taken into account when selecting (in practice) or analyzing (in research) support procedures.

## References

- De Koning, B. B., Tabbers, H. K., Rikers, R., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? *Learning and Instruction*, 20, 111–122.
- Gopher, D., Weil, M., & Siegel, D. (1989). Practice Under Changing Priorities: An approach to the training of complex skills. *Acta Psychologica*, 71, 147-177.
- Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modelling examples. *Instructional Science*.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, 14, 225-244.
- Lowe, R., & Schnotz, W. (Eds.). (2008). *Learning with animation: Research implications for design*. Cambridge, New York: Cambridge Univ.

# Does Relevance Matter in Comprehending Scientific Conflicts from Multiple Documents? Evidence from Online and Offline-Data

Marc Stadtler, Lisa Scharrer, Rainer Bromme  
University of Münster  
Fliehdnerstraße 21, 48149 Münster, Germany  
md.comprehension.lab@uni-muenster.de

**Abstract.** The goal of our study was firstly to examine whether topical relevance of conflicting information would facilitate memory for conflicts when reading multiple documents. Secondly, we sought to discriminate between two competing explanations for the assumed memory facilitation effect using eye-tracking methodology. According to the encoding-hypothesis, the relevance effect should manifest itself in an early stage of conflict processing affecting conflict detection. The elaboration-hypothesis, in contrast, assumes that the memory advantage is rather caused by differences in regulatory behaviour executed in reaction to conflict detection. Results reveal a strong effect of relevance on memory for conflicts. Fixation-time analyses support for the elaboration-hypothesis, suggesting that relevance exerts its influence late in the process of comprehending conflicting information by primarily affecting the regulation phase.

**Keywords:** multiple documents comprehension, digital reading, science understanding, eye-tracking

## Introduction

The influence of reading goals on relevance-based text processing is a well documented phenomenon in text comprehension (e.g., McCrudden & Schraw, 2007). Readers process information that is relevant against the background of their reading goal more intensely and recall it better after reading (Kaakinen, Hyönä & Keenan, 2002). However, one limitation of these findings is that they were mostly gained with non-conflicting single texts. Hence, it is unclear whether the pattern of results can be transferred to the understanding of scientific conflicts covered in multiple documents. We contend that this question is of considerable practical importance in the era of digital reading as many authentic reading situations involve making sense of multiple documents that are partly inconsistent (Stadtler, Scharrer & Bromme, 2011). Moreover, the question is not trivial from a theoretical stance, either. The status of being conflicting alone may already lend relevance to information, since it interferes with the goal of forming a coherent mental representation. Hence, the text-based importance of conflicts might level off the effects of externally imposed relevance instructions (McCrudden & Schraw, 2007).

We entertained two competing hypotheses about the processes which lead to the assumed relevance effect. According to the encoding-hypothesis, the relevance effect should manifest itself in an early stage of conflict processing. This hypothesis argues that readers encode information that is not relevant to their goal only in a shallow fashion (Kaakinen et al., 2002). Hence, resonance processes at the time of encountering conflicting information will be too weak to allow moment-to-moment conflict detection. The elaboration-hypothesis, in contrast, assumes equal conflict detection under both reading conditions. Differences in memory for conflicts should rather be caused by differences in regulatory behaviour executed in reaction to conflict detection. Whereas readers to whom a given conflict is irrelevant might simply ignore it, readers conceiving of the conflicting information as relevant might for instance reread the information or seek clarification in ensuing paragraphs.

## Method

### *Participants.*

A total of 64 undergraduates with little medical knowledge participated in the study. Out of these, 57 participants provided complete data (72% female; mean age = 24.00,  $SD = 5.1$ ).

### *Materials, Design and Task.*

Participants read seven hyperlinked texts on the topic cholesterol. Their task was to read in order to later provide a summary of a subset of information. The subset of information relevant to the experimental group contained three conflicting and three undisputed subtopics. The control group had to focus on non-conflicting subtopics exclusively. Conflicts consisted of two opposing claim-explanation-passages, each. One of them, for instance, focused on the question whether high levels of cholesterol would impede vision or not.

### *Covariates.*

To control for effects of general reading skill, participants completed two subtests from the ELVES measuring microstructural component processes of reading comprehension (Richter & van Holt, 2005). Moreover, we assessed students' epistemic beliefs with the EBAM (Kienhues, 2010). Only the ELVES subtest "Sentence-Verification-Task" (SVT) showed a significant positive correlation with our dependent measures and was hence included as a covariate in our analyses.

### *Dependent Variables.*

Memory for conflicting information was measured with the Intertextual Conflict Verification Task (ICVT) (Stadtler, Scharrer & Bromme, 2011). This task comprises 48 items, which are either paraphrases of the conflicting claims or distractor items. For each item, readers indicate whether the reading materials contained any content that stood in conflict with the claim. An individual coefficient of memory for conflicting information was calculated as the difference between the proportion of stimulus phrases that were correctly identified as conflicting and the proportion of false positives. The range of this coefficient is -1 to +1, with +1 indicating perfect performance. In addition, we measured readers' ability to apply their knowledge about conflicting information in a social knowledge building task (Moskaliuk, Kimmerle & Cress, 2012). Participants had to revise an existing Wiki so that it would better reflect the incomplete status of the topic. The number of conflicts (0 to 3) that readers integrated into the Wiki was taken as the dependent measure.

Online measures of conflict processing were calculated from readers' eye-fixations on the conflicting claim-explanation passages (Kaakinen et al. 2002). We defined first-pass fixations as fixations on previously unread text segments. First-pass reinspections were defined as regressive fixations on previously read portions of a sentence before moving on to the next sentence. Finally, look-backs were defined as regressive fixations to a previously read sentence. To control for sentence length, time parameters were qualified by number of characters. Whether readers detected the conflicts was assessed by comparing fixation times for conflicting target sentences with uncritical control sentences that were matched for relevance, familiarity and syntactic structure. Longer fixation times for target claims relative to control claims are taken as indicator of conflict detection; fixation times for explanations and measures of rereading are interpreted as indicating regulatory behaviour.

## **Results**

An ANCOVA with the SVT as a covariate confirms that readers of the experimental group showed better memory for conflicting information than controls (EG:  $M = .71$ ,  $SD = .30$ ; CG:  $M = .38$ ,  $SD = .41$ ),  $F(1, 54) = 14.30$ ,  $p < .001$ ,  $\eta^2_{\text{part}} = .21$ . This finding was corroborated by the analyses of the Wiki revision task. Experimental group readers included a significantly higher number of conflicts in their revision (EG:  $M = 2.12$ ,  $SD = 1.03$ ; CG:  $M = 1.17$ ,  $SD = 1.20$ ;  $F(1, 54) = 14.27$ ,  $p < .001$ ,  $\eta^2_{\text{part}} = .21$ ).

Readers' first-pass fixations, first-pass reinspections and look-backs for target and control claims were subjected to separate mixed 2 (claim type)\* 2 (experimental condition) ANOVAs. The analysis of first-pass fixations revealed that during initial reading, target claims were fixated significantly

longer than control claims,  $F(1, 55) = 15.67$ ,  $p < .001$ ,  $\eta^2_{\text{part}} = .22$ . Neither the main effect of experimental condition nor the interaction was significant suggesting readers of both groups were equally successful in detecting the conflicts, all  $F_s(1, 54) < 2.54$ , *ns*. Moreover, the duration of first-pass reinspections was similar across experimental condition and claim type, all  $F_s(1, 55) < 1.03$ , *ns*. However, readers of the experimental group invested significantly more time looking back to claims than controls,  $F(1, 55) = 6.72$ ,  $\eta^2_{\text{part}} = .11$ . Look-back fixations were not affected by claim type, nor was there an interaction between claim type and experimental condition, all  $F_s(1, 55) < 1.77$ , *ns*.

Likewise, all three fixation time parameters for target and control explanations were subjected to separate mixed 2 (explanation type)\* 2 (experimental condition) ANOVAs. This revealed that readers in the experimental group processed explanations more intensely than controls during initial reading,  $F(1, 55) = 8.22$ ,  $\eta^2_{\text{part}} = .13$ . No effects were found when analyzing first-pass reinspections and look-backs landing on the explanations, all  $F_s(1, 55) < 1.55$ , *ns*.

## Summary and Discussion

To summarize, our results show better performance of readers for whom conflicts were relevant in terms of memory for conflicting information and its application in a social knowledge building task. This finding extends previous research by showing that the relevance effect also applies to reading conflicting scientific information covered in multiple documents (Kaakinen et al., 2002; McCrudden & Schraw, 2007). When examining how these results came about on a process level, we found support for a late relevance effect in line with the elaboration-hypothesis. Readers seemed to initially detect conflicts regardless of their topical relevance. However, only readers for whom a conflict was relevant spent a significant amount of time looking back to conflicting claims. In addition they maintained their high level of attention when reading subsequent information that provided an explanation for the conflicting claim. This may have enabled them to better memorize the conflicting information. Readers for whom conflicts were irrelevant, in contrast, reduced their level of attention immediately after encountering the conflicting claims. Further research is needed to clarify which characteristics of the explanations (e.g., their plausibility or reinstatement of the claim) actually facilitate the integration of conflicting information into readers' mental representation.

## References

- Kaakinen, J. K., Hyönä, J., & Keenan, J. (2002). Perspective effects on on-line text processing. *Discourse Processes*, 33, 159-173.
- Kienhues, D. (2010). Whom to blame - the source of information or myself? Personal epistemology and personal ability in dealing with medical information on the Internet. Berlin: Logos.
- McCrudden, M., & Schraw, G. (2007). Relevance and Goal-Focusing in Text Processing. *Educational Psychology Review*, 19, 113-139.
- Moskaliuk, J., Kimmerle, J., & Cress, U. (2012). Collaborative knowledge building with wikis: The impact of redundancy and polarity. *Computers and Education*, 58, 1049–1057.
- Richter, T. & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur Erfassung der Effizienz von Teilprozessen des Leseverstehens, *Diagnostica*, 51, 169-182.
- Stadler, M., Scharrer, L. & Bromme, R. (2011). How Reading Goals and Rhetorical Signals Influence Recipients' Recognition of Intertextual Conflicts. In L. Carlson, C. Hölscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1346 - 1351). Austin, TX: Cognitive Science Society.

# Do Prior Attitudes Influence Epistemic Cognition While Reading Conflicting Information?

Johan L. H. van Strien, Saskia Brand-Gruwel, Henny P. A. Boshuizen  
Centre for Learning Sciences and Technologies, Open University of the Netherlands  
Valkenburgerweg 177 6419 AT, The Netherlands  
Johan.vanStrien@ou.nl; Saskia.Brand-Gruwel@ou.nl; Els.Boshuizen@ou.nl

**Abstract.** Students spontaneously engage in epistemic cognition when reading conflicting scientific information. This study examines how this epistemic cognition is related to students' actual beliefs. In addition, the interplay of students' epistemic beliefs and prior attitudes when encountering conflicting and partly attitude-inconsistent information on a controversial socio-scientific is studied using think-aloud methods.

**Keywords:** Epistemic cognition, prior attitudes, conflicting information, multiple texts

## Introduction

Students are often confronted with conflicting information on controversial socio-scientific topics when searching for information on the Internet. Such topics are characterised by the existence of competing theories and points of view for which there is no simple conclusion (Kobayashi, 2009).

### *Epistemic Cognition in Action*

Epistemic cognition pertains to the processes involved in defining, acquiring, and using knowledge (Greene, Azevedo, & Torney-Putra, 2008), and includes metacognitive thinking about the nature of knowledge and justification for knowing (Mason, Boldrin, & Ariasi, 2010). Previous work has demonstrated that university students (Ferguson, Bråten, & Strømsø, 2012) as well upper level secondary school students (Mason, Ariasi, & Boldrin, 2011) spontaneously engage in epistemic cognition when reading multiple texts on controversial topics. That is, students verbalize epistemic reflections regarding the source of knowledge, justification of knowledge, simplicity and complexity of knowledge, and certainty and uncertainty of knowledge.

### *Interplay of Epistemic Beliefs and Prior Attitudes*

In current models of epistemic cognition, the willingness to adjust beliefs is emphasized. Bendixen and Rule's (2004) model, for instance, includes among other things mechanisms of change. This starts with epistemic doubt, i.e., questioning one's beliefs, and requires epistemic volition, or the willingness to adapt these beliefs. Finally, resolution strategies are necessary to overcome doubt.

Yet, learners' prior attitudes may prevent learners from giving in to epistemic doubt or to undertake actions to change their beliefs. For instance, learners with strong prior attitudes may display a disconfirmation bias, i.e., a tendency to spend more time on and allocate more cognitive effort to information that is inconsistent with their attitudes and beliefs, while quickly and uncritically accepting information that supports these attitudes (Taber & Lodge, 2006). This could mean that students may be more inclined to see themselves as source of knowledge, but also that they are more critical of the status of attitude-inconsistent knowledge claims.

### *Research Question and Hypotheses*

As prior attitudes can thus strongly bias information processing, it is interesting to know how prior attitudes and epistemic cognition interact. The following research question were addressed: How do prior attitudes towards climate change influence students' epistemic cognition when reading multiple texts containing conflicting information? In this study, participants were classified as holding either

predominantly multiplist or evaluativist epistemic beliefs. At the multiplist level, people become aware of the uncertain and subjective nature of knowledge and knowing. However, this awareness overrules any objective standards for the evaluation of information. As a consequence, all opinions are deemed equally valuable and right. At the evaluativist level, uncertainty is acknowledged, but without ignoring the importance of evaluation of knowledge claims. That is, two positions can both be right, but one can be better supported by evidence, making it more valuable (Kuhn, 1999).

It should be expected that students holding evaluativist epistemic beliefs would act according to these beliefs and come to the conclusion that their prior attitudes may not be right. However, if prior attitudes prevail, students should – irrespective of their epistemic beliefs – attach greater value to the information that is consistent with these attitudes than to information that is in line with scientific evidence and, as a consequence, more reliable. Students with evaluativist beliefs should act according to their epistemic beliefs, as could be evidenced by more epistemic reflections regarding justification of knowledge claims by authority, regarding the justification of knowledge claims by multiple sources, (Ferguson et al., 2012). Multiplists, on the other hand, are expected to verbalize more epistemic reflections regarding personal justification of knowledge claims (Ferguson et al., 2012).

## **Methods**

### *Participants*

Participants were 25 11<sup>th</sup> grade students from a Dutch secondary school for pre-university education.

### *Measures*

Prior attitudes on climate change were measured with a 12-item questionnaire ( $\alpha = .84$ ) developed for this study. Epistemic beliefs were measured using a modified version of a questionnaire previously validated among 11<sup>th</sup> grade students, which distinguishes between multiplism and evaluativism (Van Strien, Bijker, Brand-Gruwel, & Boshuizen, 2012;  $\alpha = .64$  for both scales). Utterances in think-aloud protocols were coded based on Ferguson et al. (2012), who distinguished between utterances reflecting certainty/simplicity of knowledge, justification by authority, personal justification, or justification by means of multiple sources.

### *Materials*

Participants read 16 texts on the evidence for man-made global warming. Half of the texts describes the insights from climate science as agreed on by the vast majority of climate scientists, whereas the other half challenges these views and contains information from less reliable sources.

### *Procedure*

In a first session, 98 students from 11<sup>th</sup> grade filled in the prior attitudes measure and the epistemic beliefs questionnaire. Based on the outcomes, 25 students (12 boys; 13 girls) were selected to participate in a follow-up study. Participants were divided into one of two conditions depending on their prior attitudes. One condition consisted of students with sceptic attitudes toward climate change. The second condition consisted of students with neutral prior attitudes. In each group, participants had divergent scores on evaluativism and multiplism. During the individual session participants read a number of pre-selected texts on climate change while thinking aloud, and are asked to answer a short essay question. Participants were given 30 minutes to complete the task.

## **Results and Discussion**

Preliminary analyses using the coarse-grained coding scheme showed no significant differences in the relative number of utterances regarding certainty/simplicity of knowledge ( $t(23) = 0.26, p = .799$ ),

justification by authority ( $t(23) = 0.18, p = .857$ ), personal justification ( $t(23) = -0.02, p = .981$ ), nor justification by means of multiple sources ( $t(23) = -0.51, p = .612$ ) between sceptics and students holding neutral attitudes. In general, students produced only a small proportion of utterances regarding justification of knowledge by authority ( $M = 0.072, SD = 0.060$ ), but a relatively large proportion of utterances regarding personal justification ( $M = 0.45, SD = 0.12$ ), with most participants paying only little attention to source information. These preliminary results might imply that students have difficulties with objectively and adequately evaluating information regardless of their prior attitudes. More fine-grained analyses are planned to gain a more detailed view of students' responses to conflicting, partly attitude-inconsistent information as a function of their prior attitudes and epistemic beliefs. For instance, case studies of eight students with either pronounced evaluativist or multiplist beliefs suggest that among sceptics, those holding multiplist beliefs may be less inclined to justify knowledge claims using multiple sources than those holding evaluativist beliefs.

## References

- Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist, 39*, 69-80.
- Ferguson, L. E., Bråten, I., & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction, 22*, 103-120.
- Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 43*, 142-160.
- Kobayashi, K. (2009). The influence of topic knowledge, external strategy use, and college experience on students' comprehension of controversial texts. *Learning and Individual Differences, 19*, 130-134.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*, 16-25.
- Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: spontaneous reflections about knowledge and knowing during online information searching and their influences on learning. *Learning and Instruction, 21*, 137-151.
- Mason, L., Boldrin, A., & Ariasi, N. (2010). Searching the web to learn about a controversial topic: Are students epistemically active? *Instructional Science, 38*, 607-633.
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*(3), 755-769.
- Van Strien, J. L. H., Bijker, M., Brand-Gruwel, S., & Boshuizen, H. P. A. (2012, July). Measuring sophistication of epistemic beliefs using Rasch analysis. Paper presented at the International Conference of the Learning Sciences, Sydney, Australia.

# Not Static or Dynamic Media but Navigational Affordances Promote Active Learning

*Huib Tabbers, Jacob Diepenhorst*

Erasmus University Rotterdam

Burgemeester Oudlaan 50, 3000 DR Rotterdam, The Netherlands

tabbers@fsw.eur.nl

**Abstract.** This study tests the superiority of static media over dynamic media for promoting active learning as suggested by Mayer, Hegarty, Mayer and Campbell (2005). In a replication of their first experiment, students watched either an animation or a series of static pictures depicting the formation of lightning. However, this time the navigational affordances were manipulated separately, by providing either maximal freedom in controlling the order and time spent on each fragment of the materials, or providing no learner control. Results show a replication of the original finding as well as a main effect of navigational affordances, but no main effect of dynamism. This implies that not static media but navigational affordances may be the key factor in promoting active learning in multimedia.

**Keywords:** animations, static pictures, navigational affordance, learner control

In animation research, an on-going discussion is whether dynamic visualizations are superior to their static counterparts for supporting learning. In a meta-analysis from 2007, Höffler and Leutner concluded that on average, animations had a small learning benefit over static pictures. However, a large variance existed between individual studies, and the meta-analysis did not include one well-cited piece of evidence reporting a reverse effect. In 2005, Mayer, Hegarty, Mayer and Campbell reported the results of four experiments in which they consistently found no benefits of dynamic visualizations and in most cases even a superiority of static diagrams over animations. They concluded that for short multimedia lessons, static illustrations with printed text will promote active learning more than dynamic visualizations.

Although the study by Mayer and colleagues has been very influential, its conclusion is limited by a deliberate confound in the reported comparisons. Mayer et al. explicitly wanted to compare static and dynamic visualizations in their ‘usual’ way of presenting, as paper-based static visualizations with written text and as narrated animations. This way however, not only the dynamism of the visualization was manipulated, but also the navigational affordances. In the static condition, learners could freely navigate both the visualizations and the text, whereas in the dynamic condition learners had no control on how they could study the materials. As earlier studies have shown beneficial effects of navigational freedom on learning from visualizations (e.g., Boucheix & Schneider, 2009; Tabbers & De Koeijer, 2010), the superiority of static visualizations found by Mayer et al. could well be the result of a difference in navigational affordances rather than a difference in dynamism.

To investigate this hypothesis, we tried to replicate Experiment 1 from the Mayer et al. (2005) study, but separately manipulated the dynamism and the navigational affordances of the multimedia lesson.

## Method

### *Participants & Design*

The participants were 87 students (68 female, 19 male; age  $M = 21.2$  years,  $SD = 3.0$ ), who participated either voluntarily or for course credits. The experiment was set up as a 2x2 between-subjects design with both the Dynamism of the visualization (Static vs. Dynamic) as well as the Navigational affordances of the multimedia lesson (Minimal vs. Maximal) as experimental factors. Participants were randomly assigned to one of the four conditions.



## Materials

**Multimedia lesson.** The multimedia lesson used in this study was a 140-second lesson showing the process of lightning formation adapted from Mayer et al. (2005, Exp. 1). Four versions of the lesson were developed. In the Static versions, a series of 16 static pictures was presented depicting the main steps in lightning formation, whereas in the Dynamic versions, the visualization was animated. Moreover, in the two ‘Minimal Navigation’ versions, the visualization was presented as either a slideshow (Static-Minimal version) or as a system-controlled animation (Dynamic-Minimal version), and the explanation accompanying the visualization was presented as a narration. In both Maximal versions the explanation was presented as written text. Moreover, the visualization was presented within 16 frames distributed over 4 pages that could be navigated by clicking on one of the numbered tabs on top of the page (see Figure 1). In the Static-Maximal version the frames contained the 16 pictures, whereas in the Dynamic-Maximal version, each frame contained a fragment of the animation that was playing continuously. Also, a time bar was added that showed the passing of time.

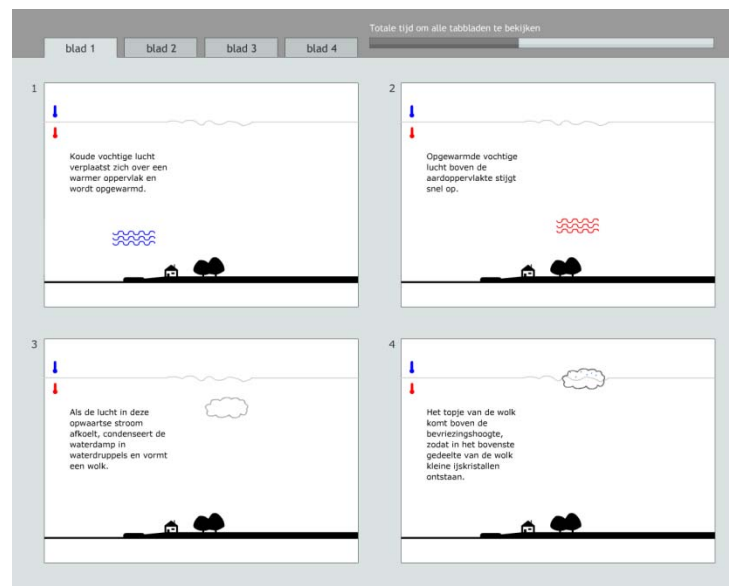


Figure 1. Screen capture of visualization with maximal navigational affordances

**Retention & Transfer.** The retention test consisted of the following open question: “Please write down an explanation of how lightning works”. A maximum of 4 minutes was allotted to write down the answer. The answers were scored by two raters on the presence of 19 major idea units (interrater correlation of .98) and the average was taken as retention score (0-19). The transfer test consisted of four open questions, like “Suppose you see clouds in the sky, but no lightning. Why not?”, with a maximum of 2.5 minutes for answering each question. Two raters scored the answers (interrater correlation of .80) and their average was taken as transfer score (0-12)

## Procedure

All participants were tested individually. They filled out a demographic questionnaire before studying the multimedia lesson. Afterwards, they made the retention test and transfer test. The total procedure took about 30 minutes.

## Results & Discussion

Table 1 shows the mean scores and standard deviations for all four conditions.

Table 1: Mean Scores on Dependent Measures.

	Retention		Transfer	
	M	SD	M	SD
Static				
Min. Navigation (n=22)	7.45	3.21	2.27	1.57
Max. Navigation (n=22)	9.00	3.81	3.41	2.77
Dynamic				
Min. Navigation (n=21)	8.24	4.96	2.29	1.55
Max. Navigation (n=22)	9.91	4.12	3.23	2.16

First we checked whether we had replicated the findings of Mayer et al. (2005). The Static-Maximal condition (Mayer: ‘paper-based text and illustrations’) and the Dynamic-Minimal condition (‘computer-based animation and narration’) did not differ significantly on retention,  $t(1,41) = 0.57$ ,  $p = .29$ , Cohen’s  $d = 0.17$ . On transfer however, there was a marginally significant difference favoring the Static-Maximal condition,  $t(1,33.3) = 1.65$ ,  $p = .05$ ,  $d = 0.50$ . This largely replicates Mayer et al. who found a small difference on retention ( $d = 0.20$ ) and a significant difference on transfer ( $d = 0.55$ ).

When analyzing the full 2x2 model however, we only found a main effect of Navigational affordance on transfer,  $F(1,83) = 5.41$ ,  $p = .02$ , partial  $\eta^2 = .06$ , favoring the Maximal over the Minimal conditions, and no main effect of Dynamism nor any interaction.

Thus, the superiority of static media over dynamic media in active learning found by Mayer may well be explained by a difference in navigational affordances provided by both representations in their experiments. This finding again shows that comparisons between static and dynamic visualizations should always be interpreted with care, and that navigational affordances are an important avenue for further multimedia learning research.

## References

- Boucheix, J. M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19, 112-127.
- Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, 17, 722-738.
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, 11, 256-265.
- Tabbers, H. K., & De Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38, 441-453.

# Benefits of a Training for Visualizing as a Learning Strategy

*Felix Wagner, David Rudolf, Tina Seufert*

Ulm University, Institute of Psychology and Education

Albert-Einstein-Allee 47, 89081 Ulm, Germany

felix.wagner@uni-ulm.de, david.rudolf@uni-ulm.de, tina.seufert@uni-ulm.de

**Abstract.** To understand learning material comprising multiple representations (i.e. text and pictures) it is crucial to build a coherent mental model of the given information. This process includes active selecting, organizing and integrating important parts of every given representation. Several studies have shown that the visualization (in forms of learner generated drawings) of the mental model can be effective when learning from text only material. Previous studies have shown that learners do not apply this strategy successfully when they are given an instruction to visualize while learning from text and pictures. Therefore we conducted a training on visualizing when learning from text and pictures. Participants were randomly assigned to either a trained or a control group. Results have shown that the trained group has generated more and better visualizations and outperformed the control group in learning outcomes and reported a lower cognitive load.

**Keywords:** learner generated visualizations, learning, multiple representations

## Introduction and Theoretical Background

Learning a certain subject in school or university mostly means dealing with multiple given representations. In order to comprehend multiple representations adequately it is crucial to build a coherent mental model. This process includes active selecting, organizing and integrating important parts of any given representation (Mayer, 2005). Recent studies on text based learning material revealed that the externalization of the mental model in forms of learner generated visualizations (logical, realistic drawings) can be an effective strategy to foster better understanding of a to-be-learned concept: for visualizing a to-be-learned concept, learners will have to mentally represent its global structure, which does not mean that the generated representation is an exact copy of the internally built mental model (Seufert, Zander, & Brünken, 2007; Van Meter, Aleksic, Schwartz, & Garner, 2006; Van Meter & Garner, 2005). Results of these studies also indicate that spatial abilities are an important factor for understanding the given subject using the strategy of visualizing (Seufert, et al., 2007). Also generating one's own representations may have a metacognitive learning effect as the learners are able to identify missing parts of the overall concept in their generated picture. Despite these expected positive effects, drawing as a strategy might increase cognitive load and hence, hinder learning (Leutner, Leopold, & Sumfleth, 2009). In a previous study we found that learners are not able to apply this strategy appropriately when simply being instructed to visualize a to-be-learned subject consisting of text and pictures. Learners generated more and better visualizations but did not benefit concerning learning outcomes and reported a higher cognitive load.

In the present study we analyzed whether a training to visualize would help learners to apply and benefit using this strategy while learning. Substantial interest is whether learners generate more and especially more comprehensive representations when they are participating in an one hour training and are asked to practice this strategy at home a week before the empirical study took part. We assume that learners in the trained group will score significantly a higher in learning outcomes (especially for higher order cognitive processes like comprehension). Learners will be able to apply this strategy more easily and report lower cognitive load scores when they received a specific training and will not be distracted or mentally overloaded by an unfamiliar strategy. Thereby, learners will be activated to build one coherent analogue representation, including relevant information of text and pictures, i.e. a mental model – hence, they should acquire deeper and more comprehensive knowledge.

## Method

In the study 51 (45 female; age  $M=21.06$ ) psychology-students took part and were randomly assigned to the training or the control group. The training group received a training one week prior to the actual study. The training was separated into two parts: participants were first given basic information about the effectiveness of visualizing and how to generate them when learning. Then they were given an assignment to practice this strategy during training and at home. When the study took part all participants were given the same learning material about a biological process within our cells which included 480 words and 7 referential pictures. Both groups received the instruction to summarize the learning content in any form on two blank sheets. Hence, no specific instruction for visualizing was given to the learners. There has been no time limit for working with the material. In addition to learners' prior knowledge specific characteristics have been assessed as control variables, such as age, sex, current task motivation and spatial abilities.

The dependent variables have been assessed after being finished working through the material and summarizing: learning outcomes including the subscales: recall, comprehension and transfer. Learning outcomes measurement consisted of 17 items including multiple-choice, free-text answers and matching tasks. Additionally the amount and quality of the constructed visualizations have been evaluated by two independent raters. The quality of the pictures has been assessed by scale rating the amount of information included within the generated pictures. Moreover cognitive load had to be rated subjectively after learning and additionally after the learning outcome test on a 5-point likert scale from 1= very low mental effort to 5 = very high mental effort (based on Paas, 1992).

## Results

Concerning the control variables, the two groups did not differ in any of the assessed variables. As recall and transfer did not have sufficient internal consistency only overall learning outcome and comprehension have been included our analyses.

For *learning outcomes* (see figure 1) we found – as expected – that the trained group outperformed the control group significantly in the overall learning outcome ( $F(1,49)=6.4$   $p<.05$ ,  $\eta^2=.12$ ) and comprehension ( $F(1,49)=4.77$   $p<.05$ ,  $\eta^2=.09$ ).

For the *amount and quality of pictures* a significant difference: the amount of pictures that have been constructed ( $F(1,49)=16.26$   $p<.01$ ,  $\eta^2=.25$ ) as well as their quality ( $F(1,49)=27.76$ ,  $p<.01$ ,  $\eta^2=.36$ ) were significantly higher in the trained group. For *cognitive load* during the learning phase the trained group reported a marginally lower value ( $F(1,49)=3.31$   $p=.075$ ,  $\eta^2=.066$ ).

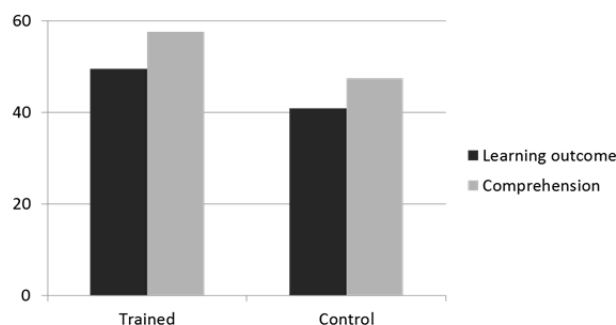


Figure 1: Percentage scores of overall learning outcome and comprehension (trained vs. control group).

## Summary and Discussion

We analyzed the effects of a training for visualizing as a strategy for learning with text and pictures. Results showed that the instruction led to a higher amount of pictures produced as well as a higher

quality. The adaption of the trained strategy also resulted in higher scores in learning outcome and a lower reported cognitive load of the trained participants.

As expected the training helped learners to apply this strategy successfully and without an increase of cognitive load. Hence, learners had sufficient cognitive resources to engage in learning processes that are inducted by the process of visualizing the mental model. As previous studies have shown that visualizing as a text learning strategy can be very successful it seems that learners need more guidance when learning with text and pictures. Only a training or more specific instructions will help learners to go beyond the given external pictorial representations and to construct more comprehensive visualizations that also comprise relevant text information. Concerning the crucial process of building a coherent mental model for learning with multiple representations, externalizing the mental model can function as an additional enhancer for the cognitive process of coherence formation.

Hence, future experimental studies will have to analyze the effectiveness of such a training for different forms of provided external representations. Additionally it will be very important to analyze what kind of generated visualizations are most supportive in activating deep comprehensive learning processes in order to specify and simplify methods and contents of future trainings. Moreover it would be interesting to also analyze metacognitive processes while learning and generating self-constructed visualizations by think aloud protocols or video analyses.

## References

- Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Comput. Hum. Behav.*, 25(2), 284-289.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge: Cambridge Univ. Press.
- Seufert, T., Zander, S., & Brünken, R. (2007). Das Generieren von Bildern als Verstehenshilfe beim Lernen aus Texten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39 (1), 33-42.
- Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, 31(2), 142-166.
- Van Meter, P., & Garner, J. (2005). The Promise and Practice of Learner-Generated Drawing: Literature Review and Synthesis. *Educational Research Review*, 17(4).

# Metaphors Activate Object Shape

Lisanne van Weelden, Joost Schilperoord, Alfons Maes  
Tilburg Center for Cognition and Communication  
Warandelaan 2, Tilburg, The Netherlands  
L.vanWeelden@uvt.nl

**Abstract.** We examined the idea that a perceptual symbol of *shape* is activated when processing a metaphoric sentence. Subjects read sentences that either did (metaphoric sentences) or did not (control sentences) invite to compare the two objects mentioned in the sentence. After reading the sentence, subjects were presented with line drawings of the two objects in question, which were either similar or dissimilar in shape. They judged whether both objects were mentioned in the preceding sentence. For the metaphoric sentences, recognition latencies were shorter for similarly shaped objects than for dissimilarly shaped objects. For the control sentences, we did not find such an effect of similarity in shape. These findings suggest that a perceptual symbol of shape is activated when processing a metaphoric sentence.

**Keywords:** Perceptual symbol theory, metaphor, object shape

Perceptual symbol theories assume that people activate perceptual representations during language comprehension (Barsalou, 1999). If we, for example, talk about a *table*, we activate the representation of an object with a flat surface and four legs. This representation is the residue of our perceptual experiences with, in this case, a table. According to Barsalou (1999), such a representation is defined by the combination of several *perceptual symbols* for different components of the referent (e.g., the color, shape, orientation, and type of wood of the table). There is an analogue relationship between these perceptual symbols and the referent. If, for example, the table is turned upside down, so too will the representation. So, any transformation the referent undergoes should cause analogous transformation in the representation.

Zwaan, Stanfield and Yaxley (2002) found evidence for this idea with their study on the effect of implied shape on mental representation. In their experiment subjects were presented with sentences like ‘*The ranger saw the eagle in the sky*’ or ‘*The ranger saw the eagle in its nest*’, which were followed by a line drawing of the object described in the sentence, in this case an eagle with outstretched wings or an eagle with folded wings. Participants recognized the pictures faster if the implied shape of the objects in the sentences matched the shape of the object in the pictures.

The purpose of the present experiment was to extend this finding. Yet where Zwaan et al. (2002) demonstrated an effect of a specific event description on the representation of the referent, we expect such an effect resulting from a comparative sentence structure. Our prediction is that inviting the reader to compare two concepts results in a mental representation of two perceptually similar concepts. In order to test this idea, we use the structural template of a metaphor: ‘*X is like Y*’. This linguistic construction is known to promote a comprehension strategy in which the two concepts are compared to each other to find out why the target (X) is similar to the base (Y) (Bowdle & Gentner, 2005). If we regard language as a set of processing instructions on how to construct a mental representation of the described situation, then this construction might result in a perceptual representation in which both concepts look like each other. That is, the comparative term *like* implies both conceptual and perceptual similarities between the two. As perceptual representations are typically schematic (Barsalou, 1999), we expect them to be similar in terms of shape.

In order to test this prediction, the present experiment examines the effect of similarity in shape on recognition latencies to two pictures of objects that were mentioned in the preceding sentence, which either did or did not invite to engage in a comparison process. An experimental group received sentences describing a metaphoric relation between two concepts based on a conceptual similarity, for example ‘*A forklift lifts heavy things, just like an elephant*’. A control group received sentences

describing a locational relation between two concepts, such as ‘*A forklift was located in front of an elephant*’. We chose for this type of sentence as control condition as the sentential structure does not invite to compare the two concepts. If readers create similarly shaped mental representations for metaphorically related objects, then recognition latencies should be faster for similarly shaped objects than for dissimilarly shaped objects. For the location sentences, however, there should be no differences in the recognition latencies between the two types of object pairs.

## Method

### Participants

Seventy-five Tilburg University undergraduates (55 women and 20 men) participated for course credit. The mean age was 21 years, ranging from 18 to 30. The data of 6 subjects were discarded because of extremely long median response latencies (>1,150 ms).

### Materials

Twenty experimental picture sets, each containing one target object and two base objects, and twenty filler picture pairs were used. The shape of the base object was depicted either similar (+Shape) or dissimilar (-Shape) to the target object, see Figure 1.

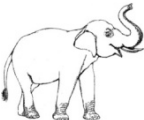

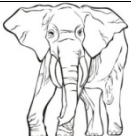
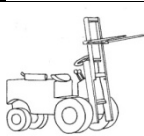
	Base objects	Target object
+Shape		
-Shape		

Figure 1. Example picture set with two picture pairs: +Shape and -Shape

Eighty Dutch sentences were created to accompany the pictures: 20 metaphoric, 20 location sentences, and 40 filler sentences. The metaphoric sentences described a conceptual correspondence between two concepts in a ‘*X has the ability to Z, just like Y*’-construction. We created the sentences either with an action verb (e.g., ‘*A motorcycle accelerates very fast, just like a leopard*’) or with a conjugation of the verb *to be* followed by a conceptual characteristic (e.g., ‘*A pawn is of relatively little value, just like a soldier*’). The location sentences described the location of the target concept in relation to the base concept. In order to create a state-of-affairs, we used sentences with verbs expressing a state of being (e.g., ‘*A bulldozer stood on top of an ant*’) and/or with action verbs in the past tense (e.g., ‘*A helicopter hovered above a dragonfly*’). The filler sentences all mentioned a base concept other than the one that was presented in the picture, and thus required a ‘no’ response.

### Procedure

During each experimental trial, subjects first saw a fixation cross in the center of the screen for 1000 ms. Subsequently, the sentence appeared. Participants pressed the ‘yes’ button when they had understood the sentence and then another fixation cross appeared in the center of the screen for 500 ms, followed by the simultaneous presentation of the two object pictures. Subjects then had to determine whether both objects were mentioned in the preceding sentence. They produced their response by pressing a key on a button panel. Immediately after their judgment, feedback indicated

whether the answer was correct, incorrect, or given too late, that is, after more than 2 seconds. Right after the feedback, the next trial started with a fixation cross on the screen. Participants were instructed that they were going to take part in a reaction time experiment and that it was important for them to make the decisions about the pictures as quickly as possible. Each subject saw 20 sentence-picture pairs, requiring 'yes' responses, and 20 filler pairs, requiring 'no' responses.

## Results

We conducted a 2 (Type of Sentence) x 2 (Shape) analysis of variance, with Type of Sentence as between subjects factor and Shape as within subjects factor, on the recognition response latencies. We analyzed response latencies of the correct responses; 3.4% of the data was excluded for this reason. In order to decrease the effects of extreme outliers, we used the median correct response times in the analyses. Table 1 displays the median response latencies.

Table 1: Median object response latencies in ms (standard deviations given in parentheses).

Type of Sentence	Shape	
	+Shape	-Shape
Metaphor (Experimental condition)*	730 (133)	793 (188)
Location (Control condition)	791 (149)	802 (145)

\* =  $p < .01$

There was an effect of Shape on response latency: Responses were faster for objects that were similar in shape than for objects that were dissimilar in shape,  $F(1, 67) = 8.94, p < .01$ . There was no effect of Type of Sentence ( $F < 1$ ). The two factors did interact,  $F(1, 67) = 4.68, p < .05$ . Post hoc analyses revealed that for the metaphoric sentences response latencies were significantly faster for objects that were similar in shape as compared to objects that were dissimilar in shape,  $F(1, 37) = 13.30, p < .01$ . Yet for the location sentences we did not find such an effect of Shape ( $F < 1$ ).

## Discussion

The result of the present study showed that the recognition of two objects that were mentioned in a metaphoric sentence – of which the sentential structure invites to compare the two objects – was influenced by the similarity in shape of the two objects. That is, recognition latencies were shorter for pictures of objects that are similar in shape than for pictures of objects that are not similar in shape. For the location sentences – which do not invite to compare the two objects –, however, we did not find such an effect of similarity in shape. These findings suggest that the structural template of a metaphor activates a mental representation of the metaphorically related objects in which the objects have a similar shape.

## References

- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193-215.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168-171.



## AUTHOR INDEX

- |                                   |                                |                           |   |
|-----------------------------------|--------------------------------|---------------------------|---|
| Ainsworth, 37, 43,<br>184         | Fabrikant, 10                  | Magner, 133               | Sanchez, 67, 79                                   |
| Airey, 76                         | Ferreira, 67                   | Maier, 136                | Scharrer, 202                                     |
| Alves, 70                         | Fidalgo, 19, 70                | Manalo, 100               | Scheiter, 3, 22, 61,<br>85, 166, 184,<br>190, 193 |
| Amadieu, 13                       | Firetto, 73                    | Mason, 169                | Schilperoord, 214                                 |
| Amelsvoort, van,<br>16            | Fredlund, 76                   | Mateos, 139               | Schnotz, 187                                      |
| Arias, 19, 70                     | Garcia Rodicio, 67,<br>79      | McElvany, 187             | Schubert, 190                                     |
| Arndt, 22, 184                    | Geiger, 82                     | Meij, H. van der,<br>145  | Schüler, 22, 61, 85,<br>190, 193                  |
| Ashraf, 25                        | Genuchten, van, 85             | Meij, J. van der,<br>145  | Schwan, 109                                       |
| Ayres, 31                         | Gerjets, 94, 97,<br>109        | Merchie, 142              | Schwartz, 11                                      |
| Baltaretu, 28                     | Glogger, 133                   | Mierlo, van, 148          | Schweppe, 196                                     |
| Baumert, 187                      | Goossens, 88                   | Mikk, 151                 | Schwonke, 133,<br>199                             |
| Bellec, 31                        | Groff, 12                      | Minguela, 118, 154        | Seufert, 82, 103,<br>211                          |
| Berney, 34                        | Hallewell, 91                  | Miras, 118                | Skuballa, 199                                     |
| Bétrancourt, 34,<br>37, 43        | Hermens, 55                    | Molinari, 34              | Solé, 121, 139, 154                               |
| Bezdan, 40                        | Hochpoechler, 187              | Münzer, 157               | Stadler, 103                                      |
| Bianco, 52                        | Holmqvist, 61                  | Nadal, 118                | Stadtler, 202                                     |
| Blanc, 9                          | Hooijdonk, van, 85             | Nardy, 52                 | Strien, van, 205                                  |
| Blat, 13                          | Horz, 187                      | Nerdel, 64, 106,<br>160   | Tabbers, 88, 148,<br>208                          |
| Bodemer, 166                      | Hoyek, 34                      | Nitz, 160, 163            | Tippett, 163                                      |
| Boshuizen, 205                    | Imhof, 94, 109,<br>184         | Nückles, 46               | Toffa, 52   |
| Boucheix, 12, 37,<br>43, 127, 172 | Kammerer, 97, 181              | Nyström, 61               | Tornatora, 169                                    |
| Brand-Gruwel, 205                 | Kaup, 109                      | Olive, 70                 | Torrance, 70                                      |
| Braun, 46                         | Keer, van, 142                 | Onghena, 112              | Trausan-Matu, 52                                  |
| Bromme, 202                       | Kester, 40                     | Paas, 148                 | Tricot, 31  |
| Camp, 88                          | Kirschner, 40                  | Pazzaglia, 193            | Uesaka, 100                                       |
| Castells, 139                     | Kiyokawa, 100                  | Pereyra, de, 178          | Ullrich, 187                                      |
| Cegarra, 13                       | Klepsch, 103                   | Philippon, 166            | Van Meter, 73                                     |
| Chevalier, 13                     | König, 103                     | Pintó Casulleras,<br>115  | Verkoeijen, 88                                    |
| Clarebout, 49                     | Kotzebue, von, 106             | Pluchino, 169             | Verschaffel, 55,<br>112                           |
| Corradi, 49                       | Kranz, 109                     | Prechtel, 64, 160         | Vogel, 82   |
| Danielson, 11                     | Kura, 100                      | Putri, 172                | Vries, de, 3, 37, 43                              |
| Dascalu, 52                       | Lamas, 139                     | Quispel, 175              | Wagner, 211                                       |
| Dessus, 52                        | Lem, 112                       | Rasch, 187                | Weelden, van, 214                                 |
| Dewolf, 55                        | Lemarié, 13                    | Renkl, 133, 199           | Westerbeek, 16                                    |
| Diepenhorst, 208                  | Linder, 76                     | Richter, 136              | Wijk, van, 28                                     |
| Dieumegard, 58                    | López Simó, 115                | Ros, 178                  | Zwaan, 88   |
| Dooren, van, 55,<br>112           | Lordán, 118, 121               | Rouet, 178                |   |
| Ehlis, 94                         | Lowe, 37, 43, 124,<br>127, 172 | Rudolf, 211               |   |
| Eitel, 61                         | Luna, 118                      | Rummer, 196               |   |
| Elen, 49                          | Macedo-Rouet,<br>178           | Salmerón, 13, 178,<br>181 |   |
| Enzingmüller, 64                  | Maes, 28, 175, 214             | Salvador Beltrán,<br>121  |   |
| Espino, 118                       | Magnat, 130                    |                           |   |

## COLLOQUE EARLI SIG 2, GRENOBLE, 2012

Le colloque biennal du *Special Interest Group 2 Comprehension of text and graphics* (Groupement d'Intérêt Spécifique 2 Compréhension de textes et de graphiques) est organisé par le Laboratoire des Sciences de l'Éducation de l'Université Pierre-Mendès-France de Grenoble du 28 au 31 août 2012. Le Groupement d'Intérêt Spécifique fait partie de l'*European Association of Research on Learning and Instruction* (Association Européenne de la Recherche en Apprentissage et Enseignement).

Le Groupement d'Intérêt Spécifique se focalise sur la façon dont l'apprentissage est influencé par les modes d'expression des contenus de l'enseignement. Historiquement, les chercheurs se sont centrés sur la compréhension de textes et d'images dans les manuels scolaires par exemple. Suite à la prolifération des technologies informatiques pour l'apprentissage (TICE, EIAH, FOAD, multimedia, e-Learning, Serious Games, langages graphiques de modélisation), il y a eu une augmentation du nombre de systèmes de représentation. Aujourd'hui, ce champ concerne l'apprentissage avec une grande diversité de représentations externes. Une « représentation externe » désigne une configuration d'inscriptions qui permet à l'élève d'interagir avec un contenu visé. Cette notion de représentation externe comprend mais n'est pas limitée à, du texte, des images, des graphiques, des diagrammes, des cartes conceptuelles, des animations, des équations, de la réalité virtuelle, de la visualisation scientifique, de l'haptique, du multimédia, de l'hypermédia et des simulations.

Le thème de l'édition 2012 est : « La mise en scène des connaissances et de l'expérience: Comment tirer parti des technologies de représentation dans l'éducation et la formation ? ». Le terme « mise en scène » est utilisé ici pour suggérer l'analogie avec la conception de l'aspect visuel des productions théâtrales et cinématographiques. En effet, les concepteurs pédagogiques, les enseignants et les formateurs, mais aussi les apprenants eux-mêmes, peuvent être considérés comme des metteurs en scène qui s'efforcent de trouver un agencement approprié des connaissances et de l'expérience à acquérir. Cette activité implique aujourd'hui d'utiliser toute une gamme de possibilités pour la représentation externe, et non pas seulement des textes et des images, comme autant de technologies sémiotiques.

Le colloque comporte trois conférenciers invités, un atelier, huit séances de communications orales et une séance de communications affichées. Nous avons reçu 75 soumissions de 176 auteurs provenant de 19 pays différents. Toutes les soumissions ont été examinées par le comité scientifique de 20 membres issus de 9 pays différents. 37 communications orales et 36 communications affichées ont été acceptées. Ce volume contient toutes les communications présentées. Le nombre élevé de contributions multi-auteurs, multi-pays montre que le groupement est en bonne santé et que le colloque biennal remplit sa fonction.

Nous remercions l'Association européenne pour la recherche en apprentissage et enseignement, le Conseil scientifique de l'Université Pierre-Mendès-France, Grenoble Alpes Métropole, la Ville de Grenoble et la Région Rhône-Alpes pour leur soutien financier. Enfin, nous tenons à remercier en particulier le comité local et les membres du Laboratoire de Sciences de l'Éducation pour l'organisation de ce colloque.

Pour plus d'informations (en anglais) :

<http://earlisig2-2012.upmf-grenoble.fr/>

<http://earli.org/>

[http://earli.org/special\\_interest\\_groups/](http://earli.org/special_interest_groups/)

[http://earli.org/special\\_interest\\_groups/comprehension](http://earli.org/special_interest_groups/comprehension)

<http://www.earli2013.org/>